

التفكير النفساني



دكتور مهدي احمد عثمان

دكتور فؤاد أبو حطب

دكتورة آمال صادق

التفكير النفساني

تأليف

دكتور فؤاد أبو حطب دكتور هيدأحمد عثمان

أستاذ علم النفس
قسم علم النفس التعليمي
كلية التربية جامعة عين شمس

أستاذ علم النفس
قسم علم النفس التعليمي
كلية التربية جامعة عين شمس

دكتورة آمال صادق

أستاذ علم النفس ووكيلة كلية التربية
جامعة حلوان

١٩٩٣

الناشر

مكتبة الأبحاث والمصيرية

١٦٥ ش محمد فريد

صدر هذا الكتاب لأول مرة بعنوان (مشكلات التقويم النفسى) عام ١٩٧٠ ثم صدر بعنوان (التقويم النفسى) كما يلى :

١٩٧٣	الطبعة الاولى
١٩٧٦	الطبعة الثانية
١٩٧٩	اعادة طبع
١٩٨٣	اعادة طبع
١٩٨٤	اعادة طبع
١٩٨٥	اعادة طبع
١٩٨٦	اعادة طبع
١٩٨٧	الطبعة الثالثة

جميع حقوق الطبع محفوظة للمؤلفين

بسم الله الرحمن الرحيم

تقديم الطبعة الثالثة

عندما صدر كتابنا الصغير المركز (مشكلات فى التقويم النفسى) فى مطلع عام ١٩٧٠ لم نكن نتوقع له أن يتطور وينمو على النحو الذى صار اليه . لقد استقبل هذا الكتاب يومئذ من دارسى علم النفس فى مصر والوطن العربى . استقبالا ابهجنا وشجعنا على أن نولى ميدان التقويم النفسى مزيدا من الاهتمام . ولم تنقضى سنوات ثلاث حتى نفذ هذا الكتاب الصغير وصدرت فى عام ١٩٧٢ الطبعة الأولى من كتاب (التقويم النفسى) فى صورته الأكثر شمولاً واتساعاً وتقدماً من الكتاب السابق . ولم تمر سنوات ثلاث أخرى الا وصدرت فى عام ١٩٧٦ الطبعة الثانية من هذا الكتاب تضيف الى الطبعة الأولى وتصحيح بعض ما جاء فيها .

وشاء - ظروفنا أن تمضى السنوات العشر الماضية دون أن تصدر طبعة جديدة من هذا الكتاب ، وكل ما نشر فى أعوام ١٩٧٩ ، ١٩٨٢ ، ١٩٨٤ ، ١٩٨٥ ، ١٩٨٦ لم يكن سوى إعادة اصدار للطبعة الثانية ، بعضه ظهر بعلمنا وعن ناشرنا ، وبعضه الآخر كان تصويراً وتزويراً للكتاب قام به ناشرون آخرون . وظاهرة تصوير الكتب وتزويرها أصبحت جزءاً خطيراً من ازمة النشر التى يتعرض لها الكتاب المصرى فى السنوات الأخيرة ، زاد من سعارها اقبال القارئ على الكتاب الجيد . وعلى الرغم من أن هذا يتضمن فى ذاته حكماً بالجودة على كتابنا الا أنه فى نفس الوقت يحمل نذر خطر اخلاقى يهدد بالشر المستطير .

كان انقطاعنا عن تجديد الكتاب طوال السنوات العشر الماضية يسبب لنا الاما مضنية . فقضايا التقويم النفسى « متعددة متجددة » كما وصفناها منذ صدر كتاب (مشكلات فى التقويم النفسى) . وقد حددنا منذ ذلك الحين مظاهر هذا التعدد والتجدد وحصرناها فى ثلاثة جوانب : اولها طبيعة الظواهر النفسية موضع التقويم ، وثانيها طبيعة أدوات التقويم النفسى ،

وثالثها طبيعة المشكلات التى تنشأ عن استخدام هذه الأدوات فى مجالات التطبيق .

والحق أن هذه الجوانب الثلاثة شهدت - منذ منتصف السبعينات - ولا تزال تشهد تطورات وتغيرات هامة يكاد يصل بعضها الى حد « الثورة العلمية » الكاملة ولعل أكثر مظاهر هذا التطور عمقا ما طرأ على النظرية السيكونومترية العامة من اتساع لآفاقها وامتداد لحدودها جعلها تتجاوز « الاختبارات المرجعة الى المعيار » الى « الاختبارات المرجعة الى المحك » ، وعلى نحو تطلب إعادة النظر فى كثير من المفاهيم الكلاسيكية لنظرية القياس النفسى .

أما أكثر مظاهر التطور حدة فى ميدان التقويم النفسى فهو ما يشهده مجال التطبيق من قضايا خلافية ظهرت على وجه الخصوص فى المجتمعات التى تعد الاختبارات النفسية جزءا من بنيتها المعرفية العامة ، وعلى وجه اخص بريطانيا والولايات المتحدة الأمريكية . وكان لهذا الجسد الحاد تضميناته الاجتماعية والثقافية والأيدولوجية .

يضاف الى ذلك ما يمكن أن نسميه « أزمة الاختبارات النفسية فى مصر والوطن العربى » . وهى أزمة معقدة متعددة الجوانب تشمل فيما تشمل ندرة الابداع فى هذا المجال السيكلوجى الهام ، والاعتماد فى كثير مما ظهر من الاختبارات النفسية على النقل المباشر أو الترجمة الحرفية أو المحرفة عن اللغات الأوربية - وخاصة اللغة الانجليزية . فاهيك عن النقص المتزايد فى اعداد المتخصصين المتمرسين فى هذا الميدان، مع زيادة مطردة فى اعداد الهواة .

ولعل أكثر ما يدعو الى الضيق حقا أن التقويم النفسى الذى يمتد عمره الزمنى فى مصر الى أكثر من خمسين عاما لم يتجاوز بعد الأروقة الأكاديمية الضيقة . وهو وضع نجد ما يشبهه فى جميع الأقطار العربية . ولعل هذا يدفعنا الى طرح تساؤل ربما يكون موضوعا لبحث علمى جدير : هل أسباب ذلك فىنا نحن السيكلوجيين الذين عجزنا عن أن نصنع قنوات اتصال جيدة مع ثقافتنا ؟ أم أنها فى مجتمعاتنا التى لا تزال تتجنب العلم سبيلا الى حل مشكلاتها ؟ أم هى فىنا وفى مجتمعاتنا معا .

كان هذا كله وغيره وراء تفكيرنا فى اصدار هذه الطبعة الجديدة من كتاب التقويم النفسى والتي تطلبت اعادة النظر فى بنيته ومحتواه جميعا .
كما تطلبت ان تشاركنا بجهودها فيها ١٠ د . امال احمد مختار صادق
فأعيد تنظيمه فى ستة أبواب . تناول الباب الأول الأسس العامة للتقويم
النفسى وجاء فى فصلين أحدهما حول ما هو التقويم النفسى ، وثانيهما يقدم
للقارئ المبتدئ مدخلا مبسوطا للاحصاء النفسى : وقد تناولنا فى الفصل
الأول معظم القضايا الخلافية المعاصرة حول الاختبارات النفسية ، أما الفصل
الثانى فهو جديد تماما وقد أضفنا الى هذه الطبعة لمحااجة القارئ اليه .

أما الباب الثانى فيعرض لمبادئ القياس النفسى ، وجاء فى أربعة
فصول تناولت على التوالى موضوعات : فن اعطاء الاختبارات ، والثبات ،
والصدق ، وتصحيح الاختبارات وتفسير الدرجات . وقد تناول هذا الباب
معظم الجديد فى النظرية السيكمترية العامة والذى شهدته طوال العقد
المنقضى .

وقد خصص الباب الثالث « لاختبارات الذكاء العام » وينقسم الى
فصلين أحدهما حول اختبارات الذكاء الفردية وعلى وجه أخص اختبارات
ستانفورد - بينيه ووكسلر ، وثانيهما حول اختبارات الذكاء الجماعية مع
تناول مفصل لهذه الاختبارات بنوعها والبحوث المصرية والعربية التى
اجريت حولها .

أما الباب الرابع فحول بطاريات الاستعدادات وينقسم أيضا الى فصلين
أحدهما حول بطاريات الاستعدادات المتعددة ، وثانيهما حول اختبارات
القدرات المتخصصة وقد عرضنا هذه الاختبارات بالتفصيل الكافى مع تناول
الجهود المصرية والعربية التى بذلت فى بناء بعض هذه الاختبارات أو فى
تقنياتها .

وأفردنا فى هذه الطبعة الباب الخامس لاختبارات التحصيل ، وجاء
هذا الباب فى فصلين تناول أولهما بناء الاختبار التحصيلى ، أما ثانيهما
فحول الاختبارات المرجعة الى المحك ، وهو فصل جديد يتناول هذا الاتجاه
الهام فى القياس التربوى .

(و)

أما الباب السادس فيتناول مقاييس الشخصية ، ويتألف من أربعة فصول تناولت على التوالى : وسائل التقرير الذاتى ، استخدام الأحكام والملاحظات المنظمة ، استخدام الأداء والأساليب الاسقاطية ، وقياس الميل .

وختمنا كل باب من أبواب الكتاب الستة بمراجعته ، ثم ختمنا الكتاب كله بقائمة مركزة تساعد المشتغل بالتقويم النفسى على تقويم أدواته التى يستعملها .

ويسعى الكتاب من خلال ذلك كله الى أن يكسب القارئ والدارس والباحث فى علم النفس بصيرة أعمق بعيدانه وبذاته وبأدواته من خلال تناول الحقائق العلمية وعرض القضايا الخلافية ، ومن خلال الإشارة الى بعض المزالق والمخاطر التى يمكن أن يقع فيها ، ومن خلال التنبيه الى بعض الأخطاء التى وقع فيها بالفعل غيره ، ولعلنا بذلك نكون قد أبلغنا اللهم فاشهد .

القاهرة :

٤ سبتمبر ١٩٨٦ م

غرة المحرم ١٤٠٧ هـ

المؤلفون

فهرس الكتاب

الصفحة	مقدمة الطبعة الثالثة
و - ج	فهرس الكتاب
ن - ز	الباب الاول : الاسس العامة
٧٤ - ١	الفصل الاول : ما هو التقويم النفسى
٥٧ - ٣	معنى التقويم (٢) - التقييم النفسى والتقويم النفسى (٤) - التقويم النفسى والقياس النفسى (٦) - مستويات القياس (٧) - طرق القياس فى علم النفس (١٦) - تعريف الاختبار (١٩) - أنواع الاختبارات (٢٤) - البرنامج التقويمى (٢٨) - دور الاختبارات فى البرنامج التقويمى (٢٢) - بعض الوظائف التربوية للتقويم النفسى (٤٤) - دور الحاسب الالىكترونى فى التقويم النفسى (٥١) - بعض المشكلات المعاصرة فى ميدان التقويم النفسى (٥٢) .

الفصل الثانى : مدخل الى الاحصاء النفسى

٥٨ - ٧١

تحليل بيانات المقاييس الاسمية (٥٩) - النسب والنسب المئوية (٥٩) - المدى المطلق (٦٠) - معامل ارتباط فاي (٦١) - تحليل بيانات مقاييس الرتبة (٦٢) - الوسطى (٦٣) - نصف المدى الربيعى (٦٤) - معامل ارتباط الرتب (٦٦) - تحليل بيانات مقاييس المسافة والنسبة (٦٨) - المتوسط الحسابى (٦٨) -

الانحراف المعياري (٦٩) - معامل الارتباط
التتابعي (٧٠) .

مراجع الباب الأول ٧٢ - ٧٣

الباب الثاني : شروط الاختبارات النفسية ٧٥ - ٢٠٣

الفصل الثالث: فن اعطاء الاختبارات ٧٧ - ٩٩

تقنين الاختبار (٧٧) - واجبات الفاحص
ومسئوليته (٧٨) - طريقة اعطاء الاختبار
(٨١) - استثارة الدوافع (٨٦) - التهيؤ لأخذ
الاختبار (٨١) - عملية الاختبار باعتبارها
علاقة اجتماعية (٩٢) - ادراك الذات وادراك
الآخر عند الفاحص والمفحوص (٩٢) - دور
اللغة في تكوين العلاقة الاجتماعية (٩٥) -
كيف يشوه الفاحصون نتائج الاختبارات (٩٧)

الفصل الرابع : ثبات الاختبارات . ١٠١ - ١٢١

معنى الثبات (١٠١) - مصادر التباين
(١٠٢) - طرق حساب الثبات (١٠٧) - اعادة
الاختبار (١٠٨) - الصور المتكافئة أو البديلة
(١١٠) - التجزئة النصفية (١١٣) - طريقة
كيورد وريتشاردسون (١١٧) - معامل الفا
لكرونبيك (١١٩) - ثبات الفاحصين
والمصححين (١٢٠) - العوامل المؤثرة في ثبات
الاختبار (١٢٢) - الخطأ المعياري للقياس
(١٢٨)

الفصل الخامس : صدق الاختبارات . ١٣٣ - ١٦٩

معنى الصدق (١٢٣) - طرق تحديد الصدق

الصفحة

- (١٢٢) - صدق المحتوى (١٢٤) - جدول
 المرافقات (١٢٧) - طرق تحديد صدق
 المحتوى (١٤٢) - الصدق المرتبط بالمحكات
 (١٤٨) - أنواع المحكات (١٥١) - بعض
 مشكلات الصدق المرتبط بالمحكات (١٥٦) -
 صدق التكوين الفرضي (١٥٧) - العوامل
 المؤثرة في الصدق (١٦٦) .

الفصل السادس: تصحيح الاختبارات

وتفسير الدرجات ١٧١ - ٢٠٠

- موضوعية الاختبار (١٧١) - تصحيح
 الاختبارات ذات الاستجابة الحرة (١٧٢) -
 تصحيح اختبارات الاختيار من متعدد (١٧٤)
 - تفسير الدرجات الخام (١٧٩) - المعايير
 (١٨٠) - المعايير الارتقائية أو النمائية
 (١٨١) - النشيطات (١٨٦) - الدرجات
 المعيارية (١٨٨) المحكات (١٩٢) - جداول
 التوقع (١٩٤) - المستويات (١٩٦) - مستوى
 نطاق المحتوى (١٩٧) - المستويات الارتقائية
 الرتبة (١٩٩)

٢٠١ - ٢٠٢ مراجع الباب الثاني

٢٠٥ - ٢٢١ الباب الثالث : اختبارات الذكاء العام

الفصل السابع : اختبارات الذكاء

٢٠٨ - ٢٧٩ الفردية

- مقياس ستانفورد - بينيه (٢٠٩) - نشأة
 المقياس وتطوره (٢١٠) - تحليل الخطأ
 المقياس (٢٢١) وصف المقياس (٢٢٤) -

تطبيق المقياس (٢٢٦) - تصحيح المقياس
 (٢٢٧) - تفسير المعايير (٢٢٤) - ثبات
 المقياس (٢٢٧) - صدق المقياس (٢٢٩) -
 مقاييس وكسلر للذكاء (٢٤٤) - مقياس
 وكسلر للذكاء الراشدين (٢٤٧) - مقياس
 وكسلر للذكاء الأطفال (٢٥٧) - المقياس
 البريطاني للقدرة (٢٦١) - بطارية كوفمان
 (٢٦٢) - قياس ذكاء الرضع وأطفال ما قبل
 المدرسة (٢٦٤) - الاختبارات العملية للذكاء
 (٢٧٥) .

الفصل الثامن : الاختبارات الجماعية

٢٨١ - ٢٢١

للذكاء

أهمية الاختبارات الجماعية (٢٨١) -
 خصائص الاختبارات الجماعية (٢٨١) -
 دور الحاسب الالكتروني (٢٨٣) - نظرية
 السمة الكامنة أو الاستجابة على المفردة
 (٢٨٤) - بطاريات اختبارات الذكاء المتعددة
 المستويات (٢٨٦) - بطاريات أطفال الحضنة
 ورياض الأطفال (٢٨٨) - بطاريات تلاميذ
 مرحلة التعليم الأساسي (٢٩٠) - اختبارات
 ذكاء طلاب المرحلة الثانوية والجامعات
 (٢٩٧) - اختبارات ذكاء طلاب الدراسات
 العليا (٣٠٠) - اختبارات ذكاء الراشدين
 المتفوقين (٣٠١) - اختبارات الذكاء المتحررة
 من أثر الثقافة (٣٠١) - اختبارات كاتل
 (٣٠٢) اختبار المصفوفات المتتابعة (٣٠٣) -
 اختبار رسم الرجل (٣٠٥) .

٣١٤ - ٣٢١

مراجع الباب الثالث

الصفحة	
٢٨٢ - ٢٢٢	الباب الرابع : مقاييس الاستعدادات . . .

٢٤٨ - ٢٢٥	الفصل التاسع : بطاريات الاستعدادات المتعددة . . .
-----------	--

- اختبارات القدرات العقلية الأولية (٢٢٦) -
- اختبارات الاستعدادات الفارقة (٢٢٢) -
- بطارية الاستعدادات العامة (٢٢٥)
- اختبارات فلاناجان لتصنيف الاستعدادات
- (٢٢٨) - بطارية اختبارات حرف المعادن
- (٢٤٢) - بطارية الاستعدادات الفارقة
- لمورسبى (٢٤٥) بطارية الاختبارات المرجعية
- العوامل (٢٤٦) - تقويم بطاريات
- الاستعدادات المتعددة (٢٤٧) .

٢٨٢ - ٢٤٩	الفصل العاشر : اختبارات القدرات المتخصصة
-----------	---

- اختبارات القدرات الحركية (٢٥٠) -
- اختبارات القدرات الميكانيكية (٢٥١) -
- اختبارات القدرة الكتابية (٢٥٤) - اختبارات
- القدرات الموسيقية (٢٥٨) - اختبارات القدرة
- فى الفنون البصرية (٢٦٤) - اختبارات
- التفكير الابتكارى (٢٦٧) - اختبارات التفكير
- الناقد (٢٧٦) - بطارية الاستعدادات الحسية
- الحركية للكفرفين (٢٧٧) - اختبارات التذكر
- (٢٧٨) - اختبار الأشكال المتضمنة (٢٧٨) .

٢٨٢ - ٢٧٩	مراجع الباب الرابع
-----------	------------------------------

الباب الخامس : اختبارات التحصيل

الفصل الحادي عشر : بناء الاختبار

٢٨٧ - ٤١٢ التحصيلي

- تحديد الأهداف التعليمية (٢٨٧) - العلاقة
- بين الأهداف التعليمية وأساليب التقييم
- (٢٩٢) - تحليل محتوى المادة الدراسية
- (٢٩٦) - اعداد جدول المواصفات (٢٩٧) -
- تحديد نوع الأسئلة (٢٩٧) - تقويم التحصيل
- في القراءة (٤٠٢) - تقويم التحصيل في
- الحساب (٤٠٦) - تقويم التحصيل في العلوم
- (٤٠٨) .

الفصل الثاني عشر : الاختبارات

٤١٢ - ٤٢٤ المرجعة الى المحك

- مقدمة (٤١٢) - مصادر المحك (٤١٤) -
- ملامح الاختبارات مرجعية المحك (٤١٥) -
- تفسير درجات الاختبار المرجع الى المحك
- (٤١٨) - الاختبارات التحصيلية المرجعة
- الى المحك (٤١٩) - مقارنة بين الاختبارات
- المرجعة الى المعيار والاختبارات المرجعة
- الى المحك (٤٢٠) .

٤٢٢ - ٤٢٤

مراجع الباب الخامس

٤٢٥ - ٥٠٢

الباب السادس : قياس الشخصية

الفصل الثالث عشر : وسائل التقرير

٤٢٧ - ٤٤٩

الذاتي

- مقدمة تاريخية (٤٢٧) - اختبار مينسوتا
- المتعدد الأوجه (٤٢٤) - اختبار مينيسوتا

الصفحة

للارشاد النفسى (٤٤٣) - نماذج أخرى ممثلة
من اختبارات الشخصية المعتمدة على وسائل
التقرير الذاتى (٤٤٥) .

الفصل الرابع عشر : الأحكام والملاحظات المنظمة . . .

٤٥١ - ٤٦٢

مقاييس التقدير وتقديرات القياس
الاجتماعى (٤٥١) - تقديرات الرؤساء
والكبار (٤٥١) - تقديرات الأقران والرسم
الاجتماعى (٤٥٥) - الملاحظة المنظمة لعينات
السلوك (٤٥٨) .

الفصل الخامس عشر : اختبارات الأداء والأساليب الاسقاطية

٤٦٣ - ٤٩٧

اختبارات الأداء ذات المواقف محددة البنية
(٤٦٥) - اختبارات المواقف غير محددة
البنية (٤٦٨) - اختبارات حل المشكلات
(٤٦٨) - الاختبارات الادراكية (٤٧٠) -
اختبار بندر جشطالت (٤٧١) - اختبار بقع
الحبر لمرورشاخ (٤٧٢) - اختبار بقع الحبر
لـبولنزمان (٤٧٧) - اختبار تفهم الموضوع
(٤٧٧) - اختبار تفهم الموضوع للأطفال
(٤٨١) اختبار تكملة الجمل (٤٨٢) - اختبار
تداعى الكلمات (٤٨٣) .

الفصل السادس عشر : قياس الميول

٤٨٥ - ٤٩٧

مقاييس الميول (٤٨٥) - اختبار سيسترونج
للميول المهنية (٤٨٧) - اختبار كيودر للميول

(ن)

صفحة

المهنية (٤٩١) - اختبار لى وثوب (٤٩٥) -
الميل الشخصية (٤٩٦) .

٥٠٢ - ٤٩٨ . . . مراجع الباب السادس

٥٠٦ - ٥٠٣ . . خاتمة : تقويم أدوات التقويم

الباب الأول الأسس العامة

الفصل الأول

ما هو التقويم النفسى ؟

معنى كلمة تقويم :

من الكلمات التى شغلت الباحثين والعامّة على حد سواء فى السنوات الأخيرة كلمة « تقويم » ، وهل لا تزال تدل على المعنى المراد منها ؟ أم الأصح أن نستخدم كلمة « تقييم » ، التى أصبحت أخف على الألسنة وأوسع فى الانتشار ؟

ولعل مصدر الخلاف هو الجذر الثلاثى للكلمة وهو « قوم » ومنه تصاغ كلمة « قيمة » تبعا للقاعدة الصرفية فى اللغة العربية التى ترى أن « الواو إذا وقعت ساكنة بعد حرف مكسور قلبت ياء لتناسب الكسرة » (١١٢ : ١٨٠ ، ١٣ ب ، ٢١٢) . * . الا أن القاعدة العامة المطردة فى الاشتقاق بالنسبة لمثل هذه الكلمات المشتقة هو العودة الى أصل الحروف فى الثلاثى . وعلى هذا ففى حالة كلمة « قيمة » نعود الى الأصل « قوم » مرة أخرى ناظرين الى الواو مرة أخرى فنقول « قوم » و « تقويما » .

ولكن لوحظ أن بعض العرب أهملوا النظر الى أصل الحرف فى الجذر الثلاثى ونظروا الى حالته الراهنة بعد إبداله فى الكلمة ، وهذا ما يسميه بعض النحاة « الاشتقاق على التوهم » ، أو ما يسميه بعض المحدثين « الاشتقاق من المشتق » ، دفعا للغموض وإزالة للبس .

ومن هذه القاعدة الخاصة أجاز مجمع اللغة العربية فى القاهرة أن

(*) يشير الرقم الأول الى رقم المرجع فى القائمة المثبتة فى نهاية كل باب من أبواب الكتاب ، بينما يشير الرقم الثانى بعد العلامة (:) الى رقم الصفحة أو الصفحات فى ذلك المرجع نفسه . وعندما توجد علامة الوصل (،) فإن ما بعدها يشير الى رقم مرجع آخر وأرقام صفحاته وهذا هو النظام المتبع فى الكتاب كله .

يقال قيمت الشيء تقييما بمعنى حددت قيمته ، وذلك للتفرقة أو ازالة للباس بين هذا المعنى وبين « قومته » بمعنى عدلته وجعلته قويا أو مستقيما .

وهكذا توجد في لغتنا المعاصرة كلمتان صحيحتان فصيحتان .
وجودهما على هذا النحو يمكن أن يحل لنا مشكلة التداخل في المعنى والخلط في الاستخدام حين نجدنا ازاء كلمتين أجنبييتين هما Valuation وفي رأينا أن أفضل ترجمة لها الآن هي « تقييم » ، والكلمة الأخرى هي Evaluation وأفضل ترجمة لها أيضا كلمتنا التقليدية المعتادة « تقويم » . فالأولى لا تتجاوز معنى تحديد القيمة أو القدر ، أما الثانية ففيها معنى التعديل والتحسين والتطوير .

هل هو تقييم نفسى أو تقويم نفسى ؟

لا شك في أن علم النفس يتضمن الكثير من عمليات « التقييم » بالمعنى الذى بيناه في الفقرة السابقة ، أو بمعنى اصدار الأحكام على « قيمة » الأشياء أو الأشخاص أو الموضوعات أو الأفكار .

فالإنسان لا يكاد يتوقف عن التقييم ، واعطاء قيمة لما يدرك . إلا أن هذا التقييم فى معظمه من النوع الذى يمكن أن نسميه « التقييم المتمركز حول الذات » ، ومعناه أن الشخص يحكم على الأشياء أو الأشخاص أو غير ذلك بقدر ما ترتبط بذاته هو . وقد يستخدم فى أحكامه هذه اعتبارات المنفعة أو الألفة أو الفائدة أو المكانة أو المركز الاجتماعى أو نقصان تهديد الذات أو سهولة الفهم والادراك . وهذه جميعا تصبغ الأحكام فى مثل هذه الحالة بصيغة ذاتية ، وقد يكون بعضها أقرب الى التخمين لأننا قد نصيب فى أحكامنا هذه وقد نخطئ بنفس الدرجة من الاحتمال .

وقد تكون أحكام الفرد ونتائج تقييمه للأشياء والأشخاص قرارات سريعة لا يسبقها فحص وتدقيق كافيين لمختلف جوانب الموضوع أو الشخص الذى يصدر عليه الحكم ، وقد تتصف أحيانا بأنها لا شعورية ، بمعنى أن الشخص منا وهو يصدر هذه الأحكام لا يعى الدلالات والقرائن والأسس التى تعتمد

عليها أحكامه ، وقد لا يستطيع « التلطف » بها مع انه يكون على درجة كبيرة من الثقة فى صحتها ، وقد تكون بالفعل كذلك ، وهى بهذه الصفة تكون اقرب الى الأحكام الحدسية .

وقد يكون التقييم النفسى مرتكزا على أحكام تتوافر لدينا عنها دلالات واضحة فى الذهن ونكون على درجة كبيرة من الوعى بها ، معتمدين فى ذلك على فهم واف وكاف للظاهرة التى نقوم بتقييمها ، وعلى تحليل دقيق لعناصرها ومكوناتها . وتكون أحكامنا فى هذه الحالة ذات طبيعة استدلالية(*)

الا أن الأمر لا يتوقف فى بعض الأحيان عند هذا الحد . ففى المجال التربوى - مثلا - تتطلب دراسة السلوك الانسانى « عملية اصدار حكم على مدى تحقيق الأهداف التربوية ودراسة الآثار التى تحدثها بعض العوامل والظروف فى تيسير الوصول الى هذه الأهداف أو تعطيله » (١١ : ٥٤٩) ومعنى ذلك أن اصدار الحكم فى مجال التربية يمكن أن يتبعه اجراء عملى يتعلق بتحسين العمالية التربوية . وقد يكون ذلك فى صورة برنامج للتعليم العلاجى مثلا ، وهذا كله بالطبع يتجاوز معنى التقييم الى معنى التقويم .

وبالمثل ، حين يجرى الباحث فى المجال الكلىنىكى دراسة على الطفل « للحكم » على مستواه العقلى ، من حيث التخلف أو المتوسط أو التفوق . ان الأمريكون أدخل فى باب « التقييم النفسى » لو لم يتجاوز مرحلة اصدار الحكم على المستوى وهو من نوع تحديد القيمة أو القدر . وهذا امر نادر فى الممارسة السيكولوجية المعتادة ، لأن الشائع أن يتبع ذلك « اجراء عملى » من نوع ما كالحاق الطفل بمدرسة للتربية الفكرية أو بفصل للمتفوقين ، أو تقديم نوع من التربية التعويضية ، أو اعادة تنظيم البيئة التى يعيش فيها وغير ذلك . وكل هذا من نوع التقويم النفسى .

ومعنى ذلك أن التقويم النفسى هو الاجراء المعتاد فى المجالات النفسية والتربوية ، وهى المجالات التى شاع فيها مفهوم « التقويم » أكثر من غيرها .

(*) التمييز بين التخمين والحدس والاستدلال يستند الى تصور نظرى لآحد مؤلفى هذا الكتاب (راجع فؤاد أبو حطب المرجع رقم ١٠) .

صحيح أن المفهوم اتسع نطاقه فى السنوات الأخيرة وامتداد الى مجالات الاقتصاد والاجتماع والسياسة والطب وغيرها الى الحد الذى دفع بعض المهتمين الى الكتابة فيما يسمى « علم التقويم Evaluation Science » ، وفى جميع الحالات يتجاوز اللفظ حدود « تحديد القيمة » الى التحسين والتطوير والتنمية . ولعل هذا أقرب الى المعقول . فالباحث فى العلوم الانسانية - بصفة عامة - لا تتشابه مهمته مع مهمة الناقد الفنى أو الأدبى التى تقتصر على « وزن » العمل الفنى والأدبى « بميزان النقد » ليحدد قيمته وجدواه ، ولا تمتد بالطبع الى اقتراح الطرق التى يمكن بها أن يتحسن العمل الفنى أو الأدبى أو يتطور ، وهو عمل أقرب الى التقييم منه الى التقويم .

والباحث فى علم النفس ليس « متفرجا » على السلوك الانسانى ولا تتحدد أهدافه فى مجرد الحكم عليه لتحديد قيمته أو قدره أو مقداره أى تقييمه وإنما تمتد فى كثير من الأحوال الى تقويم هذا السلوك بالمعنى الذى أشرنا اليه ، أى أن التقويم النفسى هو القاعدة العامة والتقييم النفسى هو الاستثناء .

ومن قبيل هذا الاستثناء ما يحدث فى مجال التربية حين تقتصر مهمة المدرسة فيما يتصل بالامتحانات على الحكم على المتعلم بالنجاح أو الفشل .

ومعنى ذلك أن مهمة الاختصاصى النفسى أقرب الى مهمة الطبيب ، التى لا تقتصر على مجرد « قراءة » ميزان حرارة الجسم أو مقياس ضغط الدم أو « مشاهدة » الرسم الكهربائى للمخ أو القلب ، لمجرد الحكم بالسواء أو المرضى ، وإنما تتجاوز ذلك الى التشخيص ثم العلاج .

التقويم النفسى والقياس النفسى :

من المفاهيم التى يحدث بينها خلط كبير فى مجال علم النفس مفهوما التقويم النفسى والقياس النفسى ، وقد حددنا فيما سبق معنى التقويم النفسى ، فماذا عن القياس النفسى ؟

يعرف تناللى (٢٧ : ٢٧) القياس فى العلم عامة وفى علم النفس خاصة بأنه « قواعد استخدام الأعداد بحيث تدل على الأشياء بطريقة تشير

الى كميات من صفة أو خاصية • ومعنى ذلك أن القياس النفسى يعتمد فى جوهره على استخدام الأعداد ، الا أنه فى صورته المحكمة يتضمن فكرة الكم والتي تعنى مقدار ما يوجد فى الشخص من خاصية معينة •

ورغم أهمية هذا التصور الكمى لخصائص الانسان الا أنه ليس شائعاً فى جميع الأحوال وبالنسبة الى جميع الخصائص ، فبعض الخصائص لايزال يستعصى على التناول الكمى المباشر وتستخدم فى وصفه « لغة كيفية » كاملة • وسواء لجأنا الى الوصف « الكمى » أو « الكيفى » للسلوك الانسانى فان ذلك يعد من قبيل البيانات التى لابد أن تتوافر للباحث النفسى قبل أن يصدر أحكامه أو يتخذ قراراته أو يقترح الاجراءات العلاجية والتصحيحية • وعلى هذا فان القياس النفسى يوفر للباحث البيانات والمعلومات الكمية التى يعتمد عليها عند القيام بعمليات التقويم النفسى ، الا أنه ليس مصدر المعلومات الوحيد ، فاذا لم تتوافر لهذا الباحث الطرق الكمية لوصف السلوك ، أو اذا استعصت الظاهرة النفسية على التناول الكمى فلا بد له من الاعتماد على الطرق الكيفية فى جمع المعلومات •

ومعنى ذلك أن التقويم أكثر عمومية من القياس ، فالتقويم النفسى يعتمد على المقاييس أو على غيرها من طرق جمع المعلومات عن السلوك الانسانى • وفى كلتا الحالتين يتضمن التقويم اصدار الأحكام القيمية واتخاذ القرارات العملية • كما أن مقادير القياس الكمية أو الصيغ اللفظية الوصفية يمكن استخدامها استخدامات مختلفة تبعاً لأهدافنا من عملية التقويم النفسى • فنتائج الاختبار التحصيلى الواحد (والتي قد تتخذ صورة قياس كمى) يمكن استخدامها فى الحكم على مدى تحقيق هدف مثل الوصول بالتلاميذ الى المستوى المعتاد من الكفاءة ، كما يمكن استخدامها فى الحكم على مدى تحقيق هدف مختلف مثل تنمية قدرة معينة ، وبالطبع تختلف فى الحالتين خطة الباحث التى يستخدمها للوصول الى مثل هذه الأحكام وما يتلوها من خطوات عملية •

مستويات القياس :

أشرنا الى أن القياس يستخدم لغة الكم ، الا أن هذا لا يعنى أن جميع

المقاييس من فئة واحدة، فقد قام علماء النفس بتحديد أنواع القياس المختلفة ودرجة ملائمة العمليات الحسابية المعروفة لكل من هذه الأنواع . ولهذا الموضوع أهميته القصوى لأننا لو لم نتناوله ببعض التعمق قد نفترض أن جميع العمليات الحسابية من جمع وطرح وضرب وقسمة يمكن أن تستخدم مع جميع نظم القياس . بل قد يستنتج البعض أن القياس يصبح مستحيلا ما لم نستخدم جميع العمليات الحسابية . ولو أخذنا ميدان القياس العقلي منذ بدايته لموجدنا وعيا مبكرا بهذه المسألة فعلى سبيل المثال نجد أنه منذ ظهور الاختبارات المبكرة في هذا المجال أشار العلماء أنه لا معنى بأن الطفل الذي نسبة ذكائه ٦٥ يكون نشاطه العقلي نصف طفل آخر نسبة ذكائه ١٥٠ ، فالأطفال يختلفون دون شك كثيرا في استجاباتهم للمراقف التي تتطلب الاستدلال المجرد ، إلا أنه لا يوجد في سلوكهم ما يبرر التعبير عن هذه الفروق في صورة معامل مقدار ٢ أو كسر مقداره $\frac{1}{2}$ أو نسبة مقدارها ٢ : ١ ، بل لا معنى مطلقا لقسمة نسبة ذكاء على نسبة ذكاء أخرى .

وحيثما تنبه علماء النفس الى هذه الخاصية التي تتوافر في الأعداد التي يستخدمونها تحققوا من أنه توجد مقاييس مشابهة خارج علم النفس . فالحرارة مثلا تقاس في العادة بمقاييس مئوية (أو فارنهایتية) فإذا انخفضت الحرارة من ٣٠ درجة مئوية أثناء النهار الى ١٥ درجة مئوية أثناء الليل لا نستنتج من ذلك أن الجو أصبح في منتصف الليل نصف دفئة أثناء الظهيرة. وذلك لأن الصفر في مقاييس الحرارة صفر اعتباطي ولا يعنى « عدم وجود حرارة على الإطلاق » . كما أن درجات الحرارة فوق الصفر لا يمكن تناولها بنفس الطريقة التي نتناول بها درجات الطول أو الوزن . ومن الواضح أن مقاييس الذكاء أقرب الى مقاييس الحرارة منها الى مقاييس الأطوال .

والواقع أن أفضل تصنيف لأنواع القياس المختلفة أو مستوياته وبخاصة في ميدان علم النفس ذلك اقترحه ستيفنس (٢٠) وفيه يقسم الطرق المختلفة لاستخدام الأعداد الى أربعة أنواع هي المقاييس الاسمية ومقاييس الرتبة والمسافة والنسبة لكل منها قواعده وحدوده وضوابطه ، ولكل منها الاجراءات الاحصائية الملائمة له . وقد طور كومبس (٢٢) هذا التصنيف وحدد العلاقات التي تربط بين المستويات المختلفة وأضاف فئات جديدة تعرضها فيما يلي :

١ - المقاييس الاسمية : هذا النوع من المقاييس يقع في أدنى مستويات القياس في ضوء مدى ملائمة الاجراءات الحسابية المعروفة له ومدى تطبيقها

فيه . وتسمى المقاييس الاسمية nominal كذلك لأن الأعداد تخصص لتعيين عناوين labels أو فئات categories ينتمى إليها الأشخاص أو الأشياء كأفراد .

ومن أمثلة استخدام الأعداد لتدل على عناوين ما يقوم به الاختصاصي الجيولوجي حين يختار عددا من عينات الصخور ويعطيها ١ ، ٢ ، ٣ ، الخ ، وفى هذه الأحوال تستخدم الأعداد كما تستخدم العناوين ، ولا يتضمن هذا بحال من الأحوال أى مطلب من التحليل الرياضى . ومن هذه الاستخدامات أيضا تلك الأرقام التى تخصص لكل لاعب من فريق كرة القدم ، أو الأرقام على علامات الطريق أو أرقام العناصر الذرية .

ويجب أن تؤكد أن أى مقياس يهتم باستخدام مقصود للأعداد . ومن هذه الاستخدامات المقصودة التسمية أو العنونة labeling . وفى هذه الأحوال لا يوجد أى غرض لاستخدام التحليل الرياضى للأعداد ، كما أن الأعداد لا تعد دالة على كميات من خصائص .

ويرتبط باستخدام الأعداد كعناوين استخدامها لتدل على مجموعات أو فئات من الأشخاص أو الأشياء . ومن أمثلة ذلك اعطاء الرمز (١) ليدل على الإناث والرمز (٢) ليدل على الذكور عند التصنيف حسب الجنس ، ويمكن أن تستخدم الأعداد فى أى تصنيف آخر للأسوياء والمضطربين ، أو حسب المهن ، أو حسب الحالة الاجتماعية ، الخ . والفرق الوحيد بين استخدام الأعداد فى التسمية واستخدامها فى التصنيف هو أننا فى حالة التصنيف يتم تجميع أكثر من وحدة بعضها مع بعض ولذلك فإن الوحدات التى تعطى نفس العدد تتشابه فى خاصية من الخصائص . وكما هو الحال بالنسبة لاستخدام الأعداد فى التسمية فإن الأعداد التى تستخدم لتدل على فئات فى نظام معين للتصنيف ليس لها أى مضمون كمى . وفى كلتا الحالتين فإن العملية الحسابية التى يمكن تطبيقها على المقاييس الاسمية هى عملية العد أو التعداد للحالات الفردية ، أما الأعداد المستخدمة ذاتها فلا يمكن أن تستخدم معها عملية الجمع أو الطرح أو الضرب أو القسمة .

ويرى كومبس أن هذا النوع من المقاييس تحكمه علاقة التساوى ،

وهذا يعنى أن أى زوج من الأشياء يجب أن ينتمى بوضوح الى نفس الفئة أو لا ينتمى اليها . وعادة ما يحكم علاقة التساوى هذه مبدأ الانتظام symmetry بمعنى أنه إذا كانت $a = b$ ، $b = c$ ، $a = c$ ، ومبدأ التعدى transitive بمعنى أنه إذا كانت $a = b$ ، $b = c$ ، $a = c$ ، وبالطبع تلعب العوامل الثقافية دورا هاما فى تحديد مثل هذه الفئات الاسمية وابتكار فئات جديدة .

٢ - مقاييس الترتيب الجزئى : يرى كومبس (٢٣ : ٤٧٤ - ٤٧٥) أنه فى الفئة الواحدة من فئات المقاييس الاسمية قد نجد ما هو أكثر من مجرد أن الوحدات التى تتألف منها الفئة متساوية فيما بينها ومختلفة عن الوحدات التى تتألف منها فئة أخرى . فقد توجد بعض العلاقات بين بعض وحدات الفئات المختلفة ، ومن هذه العلاقات أن وحدات إحدى الفئات قد تكون أكبر (أو أصغر) من وحدات فئة أخرى . فإذا كان لدينا عدد من الفئات المتكافئة وكانت هذه العلاقات موجودة بين كل فئتين منها فأننا نحصل على مقياس الترتيب الجزئى partial ordinal .

ولكى نوضح ذلك لنفرض أننا نريد قياس ما يسمى المستوى الاقتصادى - الاجتماعى ، ولنفرض أيضا على سبيل التبسيط أن هذه الخاصية تتألف من مكونين هما مستوى الدخل والمستوى التعليمى ، فأننا فى هذه الحالة إذا وجدنا أن الفرد (ا) أعلى دخلا من الفرد (ب) ، وأنه فى نفس الوقت أرقى منه فى المستوى التعليمى يمكننا أن نقول أن (ا < ب) فى المستوى الاقتصادى والاجتماعى ، وكذلك إذا كان (ب) أعلى فى مكونى الخاصية من فرد ثالث هو (ح) فإن (ب) لا يصبح فقط أعلى من (ح) فى المستوى الاقتصادى - الاجتماعى ولكن (ا) يصبح أيضا أعلى من (ح) فيه أى (ا < ح) . ومن هذا يتضح أن العلاقة هنا متعدية مثل النوع الاول من المقاييس ولكنها ليست منتظمة asymmetric لأننا فى المثال السابق نقول أنه إذا كان ا أكبر من ب فإن ب ليس أكبر من ا .

لنفرض أيضا أن لدينا شخصا رابعا هو (د) مستواد فى مكونى الخاصية اقل من (ا) وأعلى من (ج) . اننا فى هذه الحالة نستطيع القول أن $a < c$ ولكن لنفرض فى نفس الوقت أن بالرغم من أن (د) أعلى

دخلا من (ب) الا أنه اقل منه تعليما . اننا فى هذه الحالة تواجه مشكلة حقيقية لأننا لا نستطيع أن نحدد مباشرة ما اذا كان (ب < د) أو (د < ب) بالنسبة للمستوى الاقتصادى والاجتماعى . وهكذا يصبح (ب) ، (د) لا يمكن المقارنة بينهما ويتخذ مقياس المستوى الاقتصادى الاجتماعى للأفراد الأربعة (ا ، ب ، د ، د) صورة الترتيب الجزئى على النحو التالى :

ا

هـ

ب

هـ

حيث يدل على أن الشخص الأعلى له مكانة أكبر من الشخص الأدنى، وحيث لا يوجد اتصال بين (ب ، د) اللذين لا يمكن المقارنة بينهما . ويمكن الحصول على هذا النوع من المقاييس حين تتألف الخاصية أو السمة من عنصرين أو أكثر لا يرتبطان فيما بينهما بعلاقة الاضافة additive أو التعويضى Compensative .

٣ - مقاييس الرتبة : المستوى التالى من مستويات القياس هو ما يسميه ستيفنس مقاييس الرتبة ordinal scales ، ويستخدم هذا النوع حين يمكن تنظيم الوحدات فى سلسلة تمتد بين الأدنى والأعلى فى الخاصية التى نقيسها، الا اننا لا نستطيع أن نحدد بدقة الفرق بين أى رتبتين . والواقع أن الترتيب هو أكثر صور القياس بدائية (باستثناء المقاييس الاسمية) . وهذا يعنى أنه أساس لكل الصور الراقية من القياس ، ومع ذلك لا يحمل الا القليل من المعلومات عن الخاصية التى نقيسها . فحين يقوم الفاحصون فى موقف المقابلة بترتيب مجموعة من المتقدمين لشغل احدى الوظائف فانهم يستخدمون مقياس الرتبة ، كما أن المئينيات التى تستخدم كثيراً فى تفسير نتائج الاختبارات هى نوع من الترتيب وتحمل خصائص مقياس الرتبة . وفى هذا النوع من المقاييس لا يمكن استخدام العمليات الحسابية المعروفة من جمع وطرح وضرب وقسمة ، ولا يمكن استخراج متوسط الرتب أو انحرافها المعيارى ، وانما تصلح لهذه المقاييس أساليب احصائية أكثر ملاءمة . وهى من النوع الذى يسمى الاحصاء اللابارامترى ومن أشهرها معامل ارتباط الرتب .

ويمكن أن يتحول مقياس الترتيب الجزئى الى مقياس رتبة كامل اذا أمكن مساواة عناصر الخاصية بعضها ببعض . ففى مثال قياس المستوى الاقتصادى - الاجتماعى الذى أشرنا اليه يمكننا أن نساوى كل عام من التعلم بجتيه واحد من الدخل الشهرى (١٢ جنيها فى السنة) ، مثلا . وفى هذه الحالة نحصل على أساس اجرائى بسيط للمقارنة بين أى شخصين أو تصنيفهم فى فئات متساوية . ويصبح المقياس فى هذه الحالة مقياسا للرتبة مادامت عناصر السمة تخضع للتحويل الخطى البسيط بحيث يحل أحد العناصر محل الآخر على أساس قاعدة أو مبدأ مفهوم .

٤ - مقياس الرتبة المترى : من الملاحظ على أنواع المقاييس الثلاثة السابقة أن عناصر المقياس هى فئات من الأشخاص أو الأشياء ، وأن العلاقات بينها هى علاقات التساوى أو « أكبر من » ، ولا يتضمن مفهوم المسافات بين الفئات . ومعنى ذلك أننا قد نلاحظ أن (أ) أكبر من (ب) وأن (ب) أكبر من (ح) إلا أننا لا نستطيع أن نحدد ما اذا كانت المسافة بين (أ) و (ب) أكبر أو أقل من المسافة بين (ب) و (ح) . ولهذا يرى كومبس (٢٣) أن المقاييس السابقة جميعا هى من قبيل المقاييس الاسمية لفقدانها خاصية هامة فى القياس عامة وهى خاصة تحديد المسافات بين فئات الأشياء أو الأشخاص .

الا أنه بالنسبة لجميع المقاييس السابقة لو أمكننا تحديد المسافات بين الفئات فأننا نحصل على مستوى أعلى من القياس ، فمثلا اذا كانت العلاقة (أكبر من) تعود على بعض المسافات بين الأشياء المتجاورة فى مقياس رتبة ، فإن هذا المقياس يصبح مقياس رتبة مرتبا جزئيا ، أما اذا كانت هذه العلاقة تصدق على جميع المسافات يصبح المقياس مرتبا ترتيبيا كاملا ويطلق كومبس (٢٣) على المقياس فى هاتين الحالتين اسم مقياس الرتبة المترى Ordered metric scale ، ويقصد به المقياس الذى يمكن أن نجد فيه ثلاثية من الفئات

$$a < b < c$$

والمثال التالى يوضح ذلك : لنفرض أننا نريد قياس سمة « السلطة » كما تتمثل فى سلوك « الرئاسة » فأننا نستطيع ذلك فى ضوء العلاقة « أكبر من » التى أشرنا اليها فى حديثنا عن مقياس الرتبة المعتاد ، فإذا كان (أ)

يرأس (ب) ، (ح) يرأس الجماعة التى ينتمى إليها (١) فأننا نقول أن ح < أ . وهكذا عن طريق تتبع التسلسل الرئاسى يمكن اعداد مقياس رتبة .

الا أن مقياس الرتبة البسيط هذا يمكن تحويله الى مقياس رتبة مترى اذا توافرت لدينا معلومات تحدد لنا الفروق فى مقدار السلطة الذى يمارسه الشخص مقارنا بآخر . ولكى نوضح ذلك نعطى مثالا من الميدان التربوى . لنفرض أن الموجه الفنى يرأس اثنين من المدرسين الأوائل ، وأن كلا من هذين المدرسين يرأس ١٠ مدرسين عاديين ، فأننا نستنتج فى هذه الحالة أن سلطة الموجه تزيد على سلطة المدرس الأول بمقدار ٢ مدرس أول ، ٢٠ مدرسا عاديا . فإذا انتقلنا فى سلم السلطة من الموجه العادى الى الموجه الأول الذى يرأس ٢ موجهين عاديين ، يرأس كل منهم ٢ من المدرسين الأوائل اللذين يرأس كل منهما أيضا ١٠ مدرسين عاديين ، فإن سلطة الموجه الأول فى هذه الحالة تزيد على سلطة الموجه العادى بمقدار ٤٠ مدرسا عاديا (٣٠ + ١٠) ، ٤ مدرسين أوائل ، ٢ موجهين عاديين . وهكذا بالنسبة لكل مستوى من مستويات الرئاسة التربوية . ومن هذا المثال يتضح أن الموجه الأول يفوق الموجه العادى فى السلطة بمقدار أكبر مما يفوق به الموجه العادى المدرس الأول ، ونستنتج أن الزيادة فى مقدار السلطة من المدرس الأول الى الموجه العادى أقل من الزيادة من الموجه العادى الى الموجه الأول .

٥ - مقاييس المسافة : يمكن القول أن ما يميز مقياس المسافة interval عن مقياس الرتبة بمستوياته المختلفة التى اشرنا اليها أنه يسمح بتحديد مدى بعد شيئين أو شخصين بعضهما عن بعض فى الخاصية موضوع القياس ، وأن تكون هذه المسافات متساوية ويحتاج ذلك الى وضع قواعد معينة يتم الاتفاق عليها لاستخدام الأعداد حتى يمكن تحديد كم الصفة أو الخاصية فى الشيء أو الشخص . ومن أمثلة ذلك أننا يمكننا أن نحصل على مقياس مسافة للأطوال فى جماعة من الأطفال اذا لجأنا بدلا من قياس الطول مباشرة - الى قياس طول كل طفل بالنسبة لأقصر طفل فى المجموعة . وبالتالي فإن أقصر طفل فى هذه المجموعة يحصل على الدرجة صفر ، والطفل الأطول منه بخمسة سنتيمترات يحصل على الدرجة ١ ، والأطول بعشرة سنتيمترات يحصل على الدرجة ٢ ، وهكذا . والاجراء الأقرب الى الشيع فى أغلب المقاييس النفسية أن تحدد المسافات فى ضوء بعد كل طفل عن المتوسط

الحسابى لأطوال المجموعة بمقادير معيارية كما تتمثل فى الانحراف المعيارى .
لنفرض أن المتوسط ١٢٠ سم والانحراف المعيارى ٥ سم ، فافنا نستطيع فى
هذه الحالة أن نعطى + ١ للطفل الذى يزيد طوله عن هذا المتوسط بمقدار
٥ سم ، والدرجة - ١ للطفل الذى يقل طوله عن المتوسط بمقدار ٥ سم أيضا .
وهذه المسافات حول متوسط الأطوال وغيره من الخصائص يمكن حسابها
دون معرفة حقيقية بمدى بعد الأشخاص عن نقطة الصفر .

والواقع أن أغلب المقاييس النفسية والتحصيلية والعقلية من هذا
القبيل ، فنحن نقارن درجات طالبين فى الاختبار ونحدد مدى بعد كل منهما
عن المتوسط . إلا أن أهم نواحي القصور فى هذه المقاييس أنها ليس لها
صفر مطلق . فقد يحصل التلميذ على الدرجة صفر فى أحد الاختبارات
التحصيلية إلا أن ذلك لا يعنى أنه لا يوجد لديه معلومات على الإطلاق حول
موضوع الاختبار ، كما أن الطفل الذى أعطيناه صفرا فى الطول فى مثالنا
السابق لا يعنى أنه ليس له طول على الإطلاق وإنما الصفر فى هذه الأحوال
هو صفر اعتباطى .

ويمكن أن تستخدم مع مقاييس المسافة عمليات الجمع والطرح ، إلا
أن عملية القسمة بالذات لا يجوز استخدامها على الإطلاق . فلا نستطيع أن
نقسم الدرجة التى حصل عليها التلميذ (١) فى الاختبار على الدرجة التى
حصل عليها التلميذ (ب) فى نفس الاختبار ، لأن القسمة تفترض مقدما وجود
الصفر المطلق ، وبالتالي لا يمكن الحصول على النسبة (كالضعف أو
النصف) . ولتوضيح ذلك نضرب المثال التالى :

نفرض أننا طبقنا اختبارا على شخصين فحصل الأول على الدرجة ٨٠
وحصل الثانى على الدرجة ٤٠ ، ولنفرض أن الباحث الذى أعد هذا الاختبار
ضمنه بالمصدفة ١٠ مفردات أخرى يسهل على كل من هذين الشخصين الإجابة
عليها إجابة صحيحة ، وفى هذه الحالة تصبح درجة الشخص الأول ٩٠
درجة والشخص الثانى ٥٠ وفى هذه الحالة يظل الفرق بين الدرجتين ثابتا ،
أى ٤٠ ، فى الحالتين ، ولكن معامل الدرجتين لن يكون متساويا ، فبدلا من
أن يكون ٢ (أى ٨٠ ÷ ٤٠) فى الحالة الأولى يصبح ١٫٨ (أى ٩٠ ÷ ٥٠)
فى الحالة الثانية . وهكذا فافنا فى الاختبار الواحد لا توجد لدينا طريقة

لايجاد ما اذا كانت معلومات احد الأشخاص ضعف معلومات شخص آخر أو نصفها أو ثلاثة أمثالها . ولكننا حين نفترض أن كل مفردة مؤشر للمعلومات يتساوى في جودته مع أى مفردة أخرى ، فأننا لا نخرق مبادئ الرياضيات أو المنطق حين نطرح درجة الشخص الأول من درجة الشخص الثانى أو حين نجمع هذه الدرجات ونستخرج متوسطها وانحرافها المعيارى . والواقع أن الطرق المختلفة التى تستخدم فى ترجمة الدرجات الخام فى الاختبارات الى الأنواع المختلفة من المعايير - وهى الطرق التى سنشير إليها فيما بعد - تعتمد فى جوهرها على عمليتى الجمع والطرح .

٦ - مقاييس النسبة : تعد مقاييس النسبة ratio scales أعلى مستويات القياس وهى تقاس خاصية معينة فى الشيء بوحدة أو مقدار معيارى من هذه الخاصية بهدف معرفة عدد الوحدات المعيارية من هذه الخاصية التى توجد فيه . فمثلا عندما نقيس طول الحجرة بهدف عد الأمتار أو السنتيمترات (وحدات القياس) التى توجد وتكرر فى هذا الطول ، فأننا نستخدم هذا النوع من المقاييس . وتصلح مع مقاييس النسبة كل العمليات الحسابية من جمع وطرح وضرب وقسمة ومع صلاحية هذه العمليات يمكن استخدام الرياضيات العليا كالجبر والهندسة التحليلية والتفاضل والتكامل وكل الطرق الاحصائية المتقنة . وتتوافر فى مقاييس النسبة جميع خصائص مقاييس المسافة بالإضافة الى الصفر المطلق . ونحن نألف هذا النوع من المقاييس أكثر من غيره لأن جميع الأبعاد الفيزيائية المعروفة كالطول والوزن والحجم يمكن قياسها بهذه الطريقة . ولذلك نستطيع القول أن الشخص الذى طوله ١٨٠ سم ضعف الشخص الذى طوله ٩٠ سم . والواقع أن تسمية هذا النوع عن المقاييس باسم مقاييس النسبة جاءت من قابلية هذه المقاييس للقسمة أو التعبير عنها فى صورة نسبة .

وهذا النوع من المقاييس لا يوجد الا قليلا فى علم النفس ، ولا تتوافر المقاييس القليلة من هذا النوع الا حين نقيس الخصائص العقلية أو الوجدانية بواحدات فيزيائية ، كأن نقيس زمن الرجوع أو التعلم بوحداث زمنية (كالثانية أو أجزاء الثانية) .

طرق القياس فى علم النفس :

يهتم القياس فى العلم دائما بجانب معين من جوانب الأشياء ، أو خاصية من خصائصها . وعلى وجه التحديد نستطيع القول أننا لا نقيس الأشياء وإنما نقيس خصائصها ، ويصدق هذا على الانسان ، فنحن لا نقيس الطفل مثلا وإنما نقيس ذكاءه أو اجتماعيته . ولهذا التمييز أهميته من ناحيتين : أحدهما أن القياس يتطلب عملية تجريد . فالخاصية تدل على علاقة بين الأشياء فى بعد معين : كالوزن أو الذكاء ولا تختلط بالجوانب (أو الأبعاد) الأخرى للأشياء وهذه المسألة رغم وضوحها للقارئ المتخصص إلا أنها لا تبدو كذلك لغير المتخصصين ، فكثيرا ما نجد خلطا بين خاصية معينة للأشياء وكل الخصائص التى يمكن ملاحظتها فيها ، والفضل فى تجريد خاصية معينة يجعل مفاهيم القياس صعبة الفهم ، ومن ذلك مثلا يصعب على كثير من الناس أن يفهم أن الجانح والسوى يمكن أن يتوافر لهما نفس المستوى من الذكاء (كما يقاس باختبارات الذكاء) .

والسبب الثانى للتأكيد على أن القياس يتناول خاصية معينة هو أن ذلك يجبرنا على ضرورة دراسة طبيعة الخاصية قبل محاولة قياسها لأن من المحتمل ألا توجد هذه الخاصية . فليس من الضرورى أن جميع العبارات التى نستخدمها فى وصف الناس تقابلها خصائص قابلة للقياس .

ويوجد احتمال آخر أن المقياس قد يتناول خليطاً من الخصائص لا خاصية واحدة . وهذا ما نجده كثيرا فى استخبارات « التوافق » أو قوائم مشكلات الشباب والتى تتضمن مفردات ترتبط بعدد من الخصائص المنفصلة . ورغم أن هذه المقاييس المختلطة لها ما يبررها من الوجهة العملية والتطبيقية إلا أنها لا تتضمن إلا قليلا بمنطق القياس فى العلم .

والخصائص التى تخضع للقياس فى العلم قد لا تقاس مباشرة بوحدات معيارية منها ، فالحرارة تقاس بمقياس مترى ، وكذلك الوزن ، كما أن معظم هذه الخصائص من نوع التكوينات الفرضية *hypothetical constructs* التى لا تلاحظ هى فى ذاتها مباشرة ، أى أنها لا تخضع للادراك الحسى المباشر . وكثير من المفاهيم الكبرى فى العلم من هذا القبيل ، فالمغناطيسية

والنشاط الذرى وانتقال الحرارة والذكاء من نوع التكوينات الفرضية او التجريدات المعرفية . وبالطبع تحتاج دراستنا لهذه الظواهر وقياسها توافق قدر من الوقائع التى يمكن ملاحظتها او المؤشرات الخاصة بها . واذا كان علماء الفيزياء يعتمدون فى دراستهم لظواهر المغناطيسية والكهرباء والحرارة على مؤشرات مثل تغيير اتجاه ابرة البوصلة او نشاط عداد جايجر او قراءة الترمومتر ، فان علماء النفس فى دراستهم وقياسهم لتكويناتهم الفرضية كالذكاء ، والقدرة الميكانيكية والميل العلمى والاتجاه نحو العمل اليدوى وغيرها يعتمدون على مؤشرات على هذه الخصائص النفسية تنتمى الى عالم الواقع ويمكن ملاحظتها ملاحظة موضوعية مباشرة هى اساليب الاداء performance والتى تتخذ احدى صور ثلاث :

(١) الاداء اللغوى كما يتمثل فى النطق والتحدث والتلفظ شفاها وقد يمتد الى وسائل الاتصال غير اللفظى كالايماءات والاشارات وغيرها .

(ب) الاداء الحركى كما يتمثل فى نشاط أعضاء الحركة مباشرة كاستخدام الجسم أو الأيدى أو الأصابع أو الأقدام .

(ح) الاداء الفسيولوجى كما يتمثل فى نشاط الأجهزة الجسمية المختلفة كالنشاط الهرمونى أو نشاط القلب أو نشاط المخ .

ويرتبط مفهوم القياس بالعد . فان ما يفعله المرء فى أى مقياس هو عد الوحدات المتشابهة . فيقاس الطول بعدد الوحدات المتساوية من الأمتار (السنتيمترات مثلا) ، ويقاس الوزن بعدد الوحدات المتساوية فى صورة جرامات (مقدار جذب الأرض لسنتيمتر مكعب من الماء النقى) . ولا شك ان الخصائص الفيزيائية لبعض أدوات القياس لها فائدتها فى أنها تقوم بعملية العد ذاتها بدلا من أن يقوم بها الملاحظ بنفسه . وترتبط المقاييس النفسية أيضا ارتباطا وثيقا بالعد . وبالتالي فان الكرونسكوب يفيد كعداد لأجزاء الثانية فى قياس زمن الرجوع . وبعد عد أزواج الكلمات التى يتذكرها المفحوص فى تجربة عن تعلم الارتباطات الثنائية يصل الفاحص الى تحديد مقدار ما تم تعلمه . وبالمثل فالتنا بعد عد الأسئلة التى اجاب عليها المفحوص فى اختبار معين نصل الى قياس لقدرته .

(التقويم النفسى)

ويعد أداء الفرد بعد عرض قائمة الارتباطات الثنائية ، أو حله للسؤال في الاختبار مؤشرا على الخاصية النفسية موضع القياس .

ويستخدم علماء النفس عددا من الطرق في قياس الخصائص النفسية توضح لنا كلها مقدار السمة أو الخاصية في ضوء مؤشراتها ، وأهم هذه الطرق ما يأتي :

١ - تكرار أو احتمال حدوث مؤشر السمة أو الخاصية : يمكن القول أن عدد الاستجابات « الصحيحة » في اختبار للقدرة مثال لما نسميه تكرار حدوث مؤشر السمة ، لأن كل مفردة من مفردات الاختبار هي « فرصة » للدلالة على ما إذا كانت استجابة المفحوص تدفعه الى الطرف الاعلى من المقياس أو تهبط به الى طرفه الأدنى أو تضعه في المستوى المتوسط .

فمثلا إذا أجاب المفحوص اجابة صحيحة على ٦٠ سؤالا من ١٠٠ سؤال فإن قياس السمة بدلالة تكرار حدوث مؤشرها يمكن أن يتحول في هذه الحالة الى احتمال في صورة نسبة مئوية فنقول أن احتمال حدوث مؤشر السمة في هذه الحالة هو ٦٠٪ .

فاذا كان العمل الذي يؤديه المفحوص يتألف من وححدات (أسئلة) كثيرة متساوية في الصعوبة أو السهولة فيمكن ان تقاس السمة بعدد الاستجابات الصحيحة (سواء كانت في صورة تكرار أو احتمال) التي يؤديها المفحوص في وقت محدد ، وفي هذا نحن نقيس السرعة ، وقد تقاس السمة بعد الاستجابات الخاطئة التي تصدر عن المفحوص في وقت محدد وفي هذا يصبح الباحث مهتما بقياس الدقة .

٢ - شدة حدوث مؤشر السمة أو الخاصية : وتتمثل طريقة الشدة أو الحدة intensity في اختبارات القدرة في مستويات صعوبة المفردات التي يستطيع المفحوص الاجابة عليها ، وتتمثل في النشاط العضلي بمقدار الطاقة أو الجهد المبذول كما يقاس بسعة الاستجابة magnitude أو ارتفاع وعلو الاستجابة altitude أو قوة الاستجابة strength ، وفي بعض صور النشاط الأخرى فإن المكونات الانفعالية مثل ضغط الدم أو معدل

النبض أو توصيل الجلد للتيار الكهربائي أو التوتر العضلي تدل على الشدة .
وتدل شدة الاستجابة بالقبسول أو الرفض للقضايا الخلافية على قوة
الاتجاهات .

وتستخدم هذه الطريقة في القياس حين يكون على الباحث المبالغة في
تأكيد القوة . ويمكن أن تستخدم كمؤشر للسمة حين يحصل الفاحص على
مقياس يتألف من أسئلة مرتبة حسب الصعوبة . ويصبح السؤال في هذه
الحالة ما هو الحد الذي يصل اليه المفحوص ولا يتعداه ؟ ويمكن إعطاء
مثال واضح من ميدان الألعاب الرياضية وخاصة في قفز الحواجز حيث
يرفع الحاجز تدريجيا حتى يصل اللاعب الى الحد الذي لا يستطيع اجتيازه ،
ويمكن الحصول بهذه الطريقة على مقياس للأداء . ويصدق هذا على مقياس
بينيه للذكاء الذي يتألف من أسئلة متدرجة في الصعوبة ، ويتطلب هذا جهدا
كبيرا في تحديد هذه المستويات . وقد تعطى للأسئلة درجات متناسبة مع
صعوبتها بالنسبة للمفحوصين وتصبح درجة المفحوص هي مجموع هذه
الدرجات الموزونة للأسئلة التي يجب عليها اجابة صحيحة أو خاطئة (حسب
نظام وزن الدرجات) وقد يعطى للمفحوص درجة رتبة لأصعب أو أسهل
سؤال أجاب عليه .

٢ - مدى حدوث مؤشر السمة أو الخاصية : وهذه الطريقة ليست
واضحة أو شائعة الاستخدام كالطريقتين السابقتين وتتطلب جوهرها تحديد
عينة متنوعة من الأسئلة يتألف منها الاختبار ويستخدم في قياس السمة تنوع
الأسئلة التي أجاب عليها المفحوص كما يتحدد مثلا بعدد فئات الاستجابة .

تعريف الاختبار :

تعرف أنستازي (١٨) الاختبار النفسي أنه « مقياس موضوعي مقنن
لعينة من السلوك » . ويعرفه كروتباك (٢٤) بأنه طريقة منظمة لمقارنة سلوك
شخصين أو أكثر . وقد ناقش أحد المؤلفين هذين التعريفين في موضع
آخر (١٠) وتوصل منهما الى تعريف أكثر شمولا . ونبدأ بمسألة « عينة من
السلوك » التي وردت في تعريف أنستازي فنقول ان الاختبارات النفسية
تتشابه مع الاختبارات التي تستخدم في العلوم الأخرى ، بشرط أن تتم

الملاحظات على عينة صغيرة جيدة الانتقاء من سلوك الفرد . وفى هذا الصدد فإن المتخصص فى علم النفس يحذو حذو المتخصص فى علم الكيمياء مثلا حين يختبر دم المريض أو ماء الشرب فيقوم بتحليل عينة أو أكثر منه . فإذا أراد السيكولوجى أن يختبر المحصول اللفظى للطفل ، أو قدرة موظف الحسابات على أداء العمليات الحسابية أو مدى التأزر بين العين واليد عند الطيار ، فإنه يفحص أداء هؤلاء فى « عينة » ممثلة من الكلمات أو العمليات الحسابية أو الأعمال الحركية . أما مسألة شمول الاختبار السلوك موضوع الاختبار أو عدم شموله فتتوقف على وحدات العينة وطبيعتها . فالاختبار الذى يتكون من خمس مسائل فقط ، أو من أسئلة فى الضرب فقط يعد مقياسا غير جيد للمهارة العددية . واختبار المفردات اللغوية الذى لا يتضمن إلا مصطلحات الألعاب الرياضية لا يعطينا تقديرا يعول عليه للمحصول اللغوى للطفل .

وتعتمد القيمة التشخيصية diagnostic والتنبؤية predicive

للاختبار على مدى تقنيه وموضوعيته إلا أن الموضوعية والتقنين هما من شروط الاختبارات الجيدة إلى جانب شروط أخرى سنتناولها بالتفصيل فيما بعد . وتضمنها فى التعريف يخل بشرط الإيجاز اللازم . إلا أن هذه القيمة التشخيصية والتنبؤية للاختبار تعتمد فى جوهرها على مدى استخدام الاختبار كمؤشر على مجال أوسع وأكثر أهمية من السلوك . فمن النادر أن تكون عينة السلوك التى يشملها الاختبار هدف القياس النفسى فى ذاتها وإنما المهم أن تبرهن على وجود قدر من التطابق الوثيق بين عينة السلوك فى الاختبار والمجال الأكثر شمولاً الذى ينتمى إليه (بين معرفة الطفل لقائمة الكلمات فى اختبار الاستعداد اللغوى واثقانه الكلى للمفردات اللغوية مثلا كما تتمثل فى قراءاته لأحدى مجلات الأطفال) .

ويجب أن نلاحظ فى هذا الصدد أن الأمر لا يحتاج إلى وجود التشابه التام بين أسئلة الاختبار والسلوك الذى تسعى إلى تشخيصه أو التنبؤ به . وإنما الضرورى هو البرهنة على وجود تطابق تجريبي بينهما (وهذا ما يسمى الصدق) . ولذاك تتفاوت درجة التشابه بين عينة الاختبار وميدان السلوك الذى تنتمى إليه . وفى الأحوال المتطرفة قد يتطابق الاختبار تطابقا تاما مع السلوك الذى يتنبأ به أو يشخصه ، ومن ذلك اختبارات التحصيل المدرسى أو الكفاءة المهنية . وتوجد

درجة أقل من التشابه تتمثل في كثير من اختبارات الاستعدادات المهنية . وفي الطرف الآخر نجد الاختبارات التي لا ترتبط ارتباطا مباشرا بميكان السلوك ومنها الاختبارات الاسقاطية . ورغم هذه الاختلافات السطحية فان جميع هذه الاختبارات تتألف من عينات من السلوك ، ومن الواجب اثبات جدواها بالبرهان التجريبي على وجود تطابق بين أداء المفحوص في الاختبار وغير ذلك من المواقف .

الا أن الاختبار قد لا يكون بالضرورة عينة سلوك في كل الحالات . ففي ميدان الاختبارات التحصيلية يزداد الاهتمام في السنوات الأخيرة بما يسمى الاختبارات المرجعة الى المحك criterion-referenced tests وتقوم في جوهرها على مسلمة ملاحظة معطاة او جميع السلوك وليس عينة منه وخاصة حين يكون موضوع الملاحظة وحدة متكاملة من وحدات التعلم المدرسي ، وفي هذه الحالة يصنع الاختبار التحصيلي بحيث يشمل ما تم تدريسه في هذه الوحدة من محتوى وأهداف وليس عينة منهما . وفي ضوء نتائج هذه الاختبارات المرجعة الى المحك يمكن تشخيص صعوبات التعلم والتأكد من مدى الاتقان . كما يمكن اقتراح الطرق العلاجية لتحقيق أهداف التعلم التي يمكن أن يحرزها أكثر من ٩٠٪ من التلاميذ . وفي هذه الحالة تصبح العملية تقويما وليس تقويما كما هو الحال في معظم الممارسات التي تعتمد على ما يسمى الاختبارات المرجعة الى المعيار .

واستخدام كرونباك لعبارة « طريقة غير منظمة » في تعريفه للاختبار أكثر دقة من كلمة « مقياس » عند أنستازي ، لأن في قول أنستازي نوع من عدم التمييز بين المصطلح « اختبار » test والمصطلح « مقياس » measure ، فعلى الرغم من تداخل معانيهما الا انهما ليسا مترادفين .

فمن ناحية نجد أن لفظ « مقياس » أكثر عمومية لأنه يستخدم في كل ميادين البحث السيكلوجي عندما نسعى الى الحصول على أوصاف « كمية » كما هو الحال في بحوث الادراك والاحساس والحكم والمجال السيكلوفيزيائي العام . أي أن اللفظ يستخدم في الأغراض السيكلولوجية العامة ، بل وفي صميم علم النفس التجريبي . فكثيرا ما نقيس التلم أو الاستجابة أو المثير ، وتستخدم في هذه الأغراض المقاييس الفيزيائية أو النفسية .

ويطلق على المقياس لفظ اختبار فى مجال استخدامه فى ميدان علم النفس الفارق وحده . وعلى هذا فان مقاييس العتبات الفارقة أو التعلم أو الادراك يمكن أن تستخدم « كاختبارات » اذا تحول اهتمامنا بها الى ميدان الفروق الفردية . الا أن الاختيار يتكون فى العادة من عدد من الأسئلة أو المفردات التى لا تأخذ صورة مقاييس النسبة وانما قد تكون من نوع مقاييس المسافة أو الرتبة ، ومعنى ذلك أنه ليست جميع المقاييس اختبارات الا عند الاهتمام بعلم النفس الفارق . وفى هذه الحالة فقط يمكن أن يحل لفظا اختبار ومقياس كل منهما محل الآخر .

ومن ناحية أخرى ليست جميع الاختبارات مقاييس . وقد اشرنا الى معانى القياس ومستوياته وهو فى كل الأحوال يتطلب نوعا من الوصف الكمي سواء كان من نوع الكم المتصل أو الكم المنفصل . وليست جميع الاختبارات من هذا القبيل . فقد نجد بعض الاختبارات التى لا تعطى درجة للمفحوص وانما يستخدمها الأخصائى النفسى لمساعدته على الوصول الى وصف لفظى أو كیفى للمفحوص (مثل طرق الملاحظة) . ولا يتطلب الأمر فى هذه الأحوال استخدام المقاييس من أى مستوى من المستويات التى ذكرت آنفا .

ومادامت الاختبارات هى فى جوهرها أدوات الدراسة العلمية للفروق الفردية فانها تسعى الى المقارنة بين الأفراد كما يقول كرونباك فى تعريفه الا ان هذه المقارنة لا تكون بين الأفراد فى ضوء معيار norm نحسب ، وانما قد تكون أيضا داخل الأفراد فى ضوء مستوى standard أو محك criterion ، كما أن هذه المقارنة لا تكون فى عينة السلوك فقط ، كما هو الحال فى الاختبارات المرجعة الى المعيار ، وانما تشمل أيضا المقارنة فى « معظم » السلوك أو « كله » كما هو الحال فى الاختبارات المرجعة الى المحك .

والمعيار هو أساس للحكم على أداء المفحوصين والمقارنة بينهم فى ضوء أدائهم الفعلى ، ويأخذ الصيغة الكمية فى أغلب الأحوال ، ويتحدد فى ضوء الخصائص الواقعية لهذا الأداء ، ومن ذلك استخدام المتوسط الحسابى لدرجات عينة التقنين معيارا لوصف الأداء العادى فى الاختبار ، وفى ضوءه

تتحدد الأوضاع النسبية للأفراد فنقول أعلى من المتوسط أو أقل من المتوسط أو متوسط .

أما المستوى فيتشابه مع المعيار في أنه أساس للحكم على الأداء في ضوء هذا الأداء ذاته ، إلا أنه يختلف عنه في جانبين : أولهما أنه قد يأخذ الصورة الكمية أو الكيفية ، وثانيهما أنه يتحدد في ضوء ما يجب أن يكون عليه الأداء وليس ما هو عليه بالفعل . ومن هذه المستويات ما نجده في نظم الامتحانات المعتادة حين نقارن درجات التلاميذ في هذه الامتحانات بنظام النهايات الصغرى والكبرى ، أو حين تتحدد تقديرات النجاح قبلها في صورة ضعيف ومقبول وجيد وممتاز في ضوء نسب مئوية من النهاية العظمى للمادة الدراسية توضع مدمًا ولا تحسب بالطرق الاحصائية في ضوء الأداء الفعلى في الامتحانات . وهذه جميعا وسائل غير دقيقة في تحديد المستوى ، أما أفضل الطرق فتكون حين يقارن الأداء كما يحدده الاختبار بمستوى الجودة أو الاتقان أو التمكن الذي يحدده الهدف التربوى أو التعليمى أو المهنى ، ويكون تحديد هذا المستوى فى الأصل قد تم فى ضوء ما يجب أن يكون عليه الأداء .

أما المحك فهو أساس خارجى مستقل للحكم على الأداء فى الاختبار ، وقد تكون هذه المحكات كمية أو كيفية ، فمثلا لكى نحكم على نجاح برنامج تعليمى أو تدريبي فى تحقيق أهدافه يمكن مقارنة أداء التلاميذ فى الاختبارات التحصيلية المرتبطة بهذا البرنامج بمستويات الكفاية الانتاجية التى تتحدد فى الميدان الفعلى للعمل .

وهكذا توصل فؤاد أبو حطب (١٠) الى تعريف يقترحه لمفهوم الاختبار وفيه يقول :

« الاختبار النفسى هو طريقة منظمة للمقارنة بين الافراد أو داخل الفرد الواحد فى السلوك أو فى عينة منه فى ضوء معيار أو مستوى أو محك » .

انواع الاختبارات :

يمكن أن نقسم الاختبارات على أسس كثيرة منها الشكل أو الغرض أو المحتوى ويقسمها كرونباك قسمين :

١ - اختبارات الأداء الأقصى maximum performance التي تستخدم إذا كنا نريد أن نعرف إلى أي حد يستطيع الشخص أن يقرر بأداء ما إلى أقصى قدرته ، ويمكن الإشارة إليها باسم اختبارات القدرة .

٢ - أما القسم الثاني فيتضمن الاختبارات التي تهدف إلى تحديد الأداء المميز typical performance وهي تقيس ما يحتمل أن يفعله الشخص في موقف معين أو في نوع معين من المواقف ، وتدخل في هذا النطاق اختبارات الشخصية والميول .

١ - اختبارات الأداء الأقصى (أو اختبارات القدرة) :

أهم ما يميز هذه الاختبارات أن الشخص يشجع فيها ليحصل على أحسن درجة يمكنه الحصول عليها ، وتشمل اختبارات الذكاء واختبارات القدرات المختلفة سواء كانت عقلية أو حركية . ومن هذه الاختبارات مجموعة كبيرة تسمى بمقاييس القدرة العقلية العامة وهي تهدف إلى قياس تلك القدرات العقلية العامة في معظم أنماط التفكير والتعلم (وهي بهذا المعنى اختبارات ذكاء) . وهناك اختبارات القدرات الأكثر تخصصاً وهي القدرات ذات القيمة في مهام محدودة ، منها مثلاً الفهم الميكانيكي ، والاحساس بالدرجة الصوتية والمهارة الأصبعية . وهذه تسمى اختبارات القدرات المتخصصة . وهذه الاختبارات قد تستخدم منفردة أحياناً كما قد تستخدم مجتمعة ، وقد تستخدم مجموعة منها تقيس قدرات خاصة متعددة لتساعد في رسم بروفيل لقدرات الفرد .

أما اختبارات الكفاءة proficiency فتقيس القدرة على أداء مهمة أو عمل ما له أهميته في ذاته : مثل قراءة الفرنسية ، أو العزف على البيانو أو إصلاح خلل في ماكينة طائفة . وحيث أن هذه الاختبارات تقيس أداء

أشخاص تدربوا على العمل المقاس فانها تسمى احيانا اختبارات التحصيل
achievement .

وهناك اختبارات الاستعدادات aptitude وتستعمل هذه
الاختبارات للتنبؤ بالنجاح في مهنة أو في تدريب معين . فهناك مقاييس
للاستعداد الهندسي مثلاً ، أو الاستعداد الموسيقي أو الاستعداد للجبر وهكذا .
ولا تختلف هذه الاختبارات من حيث الشكل عن الأنواع الأخرى . فاختبار
الاستعداد الهندسي مثلاً قد يتضمن اقساماً تقيس القدرة العقلية العامة .
وأخرى تقيس الفهم الميكانيكي أو المكان وأقساماً تقيس التحصيل في
الرياضيات . ويسمى الاختبار على أي حال ، اختباراً تحصيلياً إذا كان
يقيس نجاح الشخص في الدراسة السابقة . ويسمى اختبار استعداد عندما
يستعمل للتنبؤ بنجاحه في عمل المستقبل .

٢ - اختبارات الأداء المميز :

لا تستعمل اختبارات الأداء المميز لمعرفة ماذا يستطيع الشخص أن يفعله
بل لمعرفة ما يعمل بالفعل وطريقة عمله (مثلاً في الاختبار لإدارة عمل ليس
المهم ما يمكن أن يعمل الشخص المرشح ولكن ما يعمل بالفعل وأسلوبه فيه) .
وهناك بعض الخصائص الشخصية . مثل البشر أو الأمانة أو سعة الأفق
لا تستطيع اختبارات القدرة أن تعرفنا بها ، على الرغم من أهميتها ، فالأداء
المميز أو السلوك الفعلي هام من غير شك إذا نحن أردنا أن نقدر الاستعداد
للنجاح في عمل ما .

وإذا كانت الدرجة العالية شيئاً مرغوباً فيه في اختبارات القدرة ، ففي
اختبارات الأداء المميز لا نستطيع أن نقول أن استجابة ما في اختبار من
اختباراتها « جيدة » ، إذ ليس هناك ما هو « جيد » أو « رديء » ، مثلاً في أن
يكون عند شخص ما ميل للهندسة . وكذلك نجد أن بعض الناس يقلب عليهم
سلوك السيطرة (وهذه صفة من صفات الشخصية) ، ونحن لا نستطيع أن
نحدد أن درجة معينة من السيطرة هي الأنسب والأفضل لأن العالم يتسع
لأناس مختلفي الأنماط والصفات .

ونستطيع القول بصفة عامة أن سلوك الشخص المميز له هو أفضل مفتاح لشخصيته ، فالعادات السلوكية لها قيمة تنبؤية في ذاتها ، فما فعله الشخص من قبل يكون أميل الى أن يعيده ويكرره في مواقف قالية فيما بعد ، ولكن يعترض معظم علماء النفس على افتراض أن عادات الشخص المنظورة هي شخصيته ، ذلك أن مواقف جديدة تنشأ دائما ، وأن أى وصف للعادات المألوفة عند شخص ما لا تدل بصورة مباشرة على ما سيفعله في موقف جديد . فربما عرف عن شاب أنه يكره البنات ثم تأتي بنت بعينها فتأسره بحبها ويدله بها . والمعالج النفسى الذى يستطيع أن ينشئ علاقة متينة وعميقة وينجح مع كثير من عملائه قد لا يستطيع أن ينشئ مثل هذه العلاقة مع شخص معين لأنه يثير عنده - أى عند المعالج - الكره والغضب مثلا . ولكن مع هذا ، ينبغي أن لا تنسينا هذه الاستثناءات أو الظروف الشاذة أن السلوك وردود الأفعال للمواقف المختلفة تعتبر حقا انعكاسات لتركيب شخصية أساسى ومتسق ، وهذا التركيب الأساسى والمتسق ينبغي استنتاجه من السلوك ، ويأمل العالم النفسى أنه عندما يفهم تركيب الشخصية هذا سوف يستطيع التنبؤ باستجابات الشخص وسلوكه حتى في المواقف الجديدة .

ويعتبر قياس الأداء المميز أمرا شاقا الى درجة بعيدة ، وقد حقق النجاح بنسب متفاوتة . وبطرق متعددة . وهذه الطرق يمكن تقسيمها الى قسمين . الأول : ملاحظات السلوك ، والثانى : وسائل التقرير الذاتى . أما ملاحظات السلوك فهي محاولات لدراسة الشخص وهو في حالة سسلوكه الطبيعى . والملاحظات يمكن أن تتم في موقف اختبار مقنن أو في موقف غير مقنن أو طبيعى . أما الملاحظة المقننة فتتطلب أن يوضع كل شخص في نفس الموقف الذى يوضع فيه الآخرون ، فيمكن أن تلاحظ « الشخصية » في موقف اختبار عقلى ، أو أثناء مناقشة جماعية . كما أنه من الممكن أن تبتكر مهام أو أعمال تتيح مجالا طيبا للملاحظة . وتسمى هذه المهام عندئذ اختبارات الأداء للشخصية .

والملاحظة المقننة تسمح بمقارنة شبه تامة لأشخاص لا يمكن أن يلاحظوا عادة في ظروف متشابهة . هذا فضلا عن أنها تكشف عن خصائص لا نشاهدها الا عرضا في حياتنا اليومية . وقد استعملت هذه الطريقة مثلا في ملاحظة ردود الفعل المميزة للحبوط وذلك بأن يتاح للشخص أن يبدأ عملا

ولكن يحال بينه وبين أن يحقق غرضه أو يصل إلى هدفه بشكل ما ويعطى رد فعله أو الطريقة التى يسلك بها فى موقف الحبوط ذاك صورة واستبصارا فى مدى ضبطه الانفعالى .

وإذا كان للملاحظة أن تظهر السلوك المميز فعلا فلا بد أن لا يعلم الشخص الخاصية أو الصفة التى يراد ملاحظتها أو تجرى ملاحظتها فعلا ، وفى موقف الحبوط المشار إليه منذ قليل من الممكن أن نذكر للشخص أو الأشخاص الذين نختبر ضبطهم الانفعالى ، أننا نقيس ذكاءهم . على أننا ينبغي أن نؤكد عند هذه النقطة أن هذا لا يجوز إلا فى مجال البحث العلمى والتجريب ، ولا بد أن يذكر للأشخاص الذين تجرى ليهم التجربة الهدف وما تم فيها بعد أن تنتهى التجربة وكذلك يجوز استخدام هذه الطريقة فى مجال الاختبار لعمل ما لمعرفة مدى توفر صفة أو صفات معينة مناسبة لعمل معين .

وكذلك يمكننا أن نحصل على معلومات عن السلوك المميز بملاحظة عينات من سلوك الشخص العادى فى نشاطاته اليومية فى الميدان كما هى ، كأن نلاحظ الأطفال فى الملعب حيث تظهر عاداتهم وشخصياتهم ، وكذلك الموظفين فى مكاتبهم ، أو العمال فى مصانعهم ، وهكذا . وطريقة الملاحظة الميدانية هذه يمكن أن تستخدم فيها أدوات تسجيل مقننة (وأعلى هذه الأدوات من حيث تقنياتها التصوير السينمائى) ، أو يستخدم فيها التسجيل والحكم الانطباعى .

أما القسم الثانى من طرق قياس الأداء المميز فهو وسائل التقرير الذاتى والأساس الذى تقوم عليه هذه الوسائل هو أنه لما كان الشخص تتاح له فرصة كبيرة ليلاحظ سلوكه الذاتى ، فإنه يستطيع إذا أراد ، أن يعطى تقريراً له قيمته عن سلوكه المميز ويساعد فعلا فى التعريف بهذه الناحية . وتستخدم الاستخبارات أو الاستبيانات questionnaires كوسيلة من وسائل التقرير الذاتى عن الأداء المميز . على أن المشكلة الرئيسية فى التقرير الذاتى ، إذا كنا سوف نفسره على أنه صورة للسلوك المميز ، هى الأمانة ، فإذا حاول الشخص أن يعطى أفضل صورة ممكنة عن نفسه بدلا من وصفها كما هى فإن الاختبار يفشل فى تحقيق غرضه ، وحتى عندما يحاول أن يكون صادقا فلن يخلو تقريره عن ذاته من خطأ تعريف أو تحيز .

ومعظم استخبارات التقرير الذاتى تعطى صورة تكاد تكون كاملة عن الشخصية ، وبعضها متخصص فى موضوعه ، فهناك مثلا استخبارات قياس عادات الدراسة ، واستخبارات الميول ، واستخبارات الاتجاهات الاجتماعية ، وهكذا . وهناك نوع آخر من استخبارات التقرير الذاتى تسمى استخبارات التكيف أو اختبارات الخلق . وهذه الأسماء تدل على الطريقة التى تفهم بها درجات الاختبار ولكنها لا تميز شكلا معيناً للاختبار .

ومن المتفق عليه بصفة عامة ألا تستعمل استخبارات الشخصية كلمة اختبار فى عناوينها ، وذلك منعا لسوء استعمال غير المدرسين على اعطاء هذه الاختبارات وتفسيرها وسوء فهم لوظيفة الاستخبارات فى التقويم النفسى .

البرنامج التقويمى :

البرنامج التقويمى هو خطة منظمة للحصول على المعلومات والبيانات الصحيحة عن الأفراد موضوع التقويم ، ولكى نوضح طبيعة البرنامج التقويمى نعطي مثالا من المجال التربوى ، حيث يعنى البرنامج التقويمى مجموعة ادوات القياس والتقويم التربوية والنفسية الشاملة والمتنوعة التى تطبقها مجموعة من المختبرين والمرشدين النفسيين والتربويين والمشرفين الاجتماعيين والمدرسين ، وفق نظام مرسوم ، يحقق حصول المؤسسة التربوية على معلومات وبيانات ترتب وتنظم بحيث يستطيع المشرفون على العملية التربوية أن يستعملوها فى اتخاذ أنسب القرارات للمتعلمين وللمؤسسة التربوية نفسها .

أى أن البرنامج التقويمى يشمل : أدوات قياس ، ومختبرين ، وخطة للاختبار والتقويم ، وبيانات منظمة ، وطرق الاستفادة من هذه البيانات ، ثم علاقات اجتماعية وتفاعل اجتماعى بين هيئة التقويم وبين سائر أعضاء المؤسسة التربوية ، وبينها وبين مؤسسات اجتماعية أخرى فى المجتمع تقتضى طبيعتها وطبيعة العملية التقويمية الاتصال بها .

ويجب أن تكون نظرتنا الى البرنامج التقويمى نظرة صحيحة فيكون العمل فيه محققا لأهدافه وأغراضه ، والنظرة الصحيحة الى البرنامج التقويمى هى التى تعتبره متكاملا مع النظام التربوى كله ، كما تعتبره جزءا من التعليم ومن التخطيط التربوى . أى أنه لا يصح أن ننظر الى البرنامج

التقويى على أنه شيء منفصل عن تيار العمل التربوى فى مؤسساته المختلفة .
أو منفصل عن عملية التعلم ، أو خارجها أو فوقها . ولن يتحقق نجاح لبرنامج
تقويى تربوى لا يقوم على هذا الأساس الكبير : التكامل مع النظام التربوى
ومع عملية التعليم ذاتها .

والمسئول عن وضع برنامج تقويى وتسييره ، فى أى مستوى من
مستويات العملية التربوية ، فى مدرسة أو معهد أو كلية أو منطقة تعليمية أو
قطر بأسره ، يمكن أن يسير وفق خطوات عامة فى بنائه للبرنامج التقويى على
أن تهديه فى خطواته نظرة صحيحة الى التقويم .

أما هذه الخطوات ، فيمكن تلخيصها فيما يلى :

١ - تحديد الأهداف : على هذا المسئول أن يبدأ بتحديد أهداف العملية
التربوية العامة التى يضعها المجتمع ويتفق عليها والتى تحددها السلطات
التربوية المسئولة التى تعبر عن أهداف هذا المجتمع وحاجاته ، ثم عليه أن
يحدد الأهداف التربوية الخاصة بالمؤسسة التى سيتولى إقامة برنامج تقويى
فيها . وبعد أن يحدد هذه الأهداف العامة والخاصة عليه أن يقوم بتعريف
القائمين على العمل التربوى فى المؤسسة بهذه الأهداف ، ويشاركهم الرأى
والناقشة فى مبرراتها وتفاصيلها ووسائل تحقيقها . ولا بأس بأن يوضع دليل
مطبوع عن البرنامج التقويى وعلاقته بهذه الأهداف يفيد منه المدرسون والآباء
ومن يعينهم الأمر .

٢ - جمع البيانات : وبعد تحديد الأهداف والتعريف بها ، على المسئول
عن التقويم أن يبدأ عملية اختيار الاختبارات التى يمكن الحصول عليها والتى
تناسب كل هدف من هذه الأهداف ، مراعى تنوع هذه الاختبارات وتمثيلها
لكل الجوانب التى يراد قياسها . ومن الضرورى مراعاة الشروط الأساسية
فى اختيار الاختبارات والتى سنتناولها بالتفصيل فيما بعد .

أما إذا لم تكن هناك اختبارات أو مقاييس مناسبة لأهداف معينة ،
فعلى القائم على البرنامج التقويى بالاشتراك مع هيئة مساعدة له من
المتخصصين فى القياس ، أن وجدت ، أو من المهتمين فى المؤسسة التربوية ،

أن يبدأ فى وضع اختبارات ووسائل قياس خاصة ومناسبة . ويمكن بطبيعة الحال الاستعانة ببعض العلماء المتخصصين فى القياس النفسى فى وضع هذه المقاييس ، وأن يراعى فى جميع الأحوال شروط بناء هذه الأدوات .

والهدف من اختيار الاختبارات أو بنائها هو استخدامها فى جميع البيانات اللازمة للتقويم ، وقد يحتاج الأمر الى بيانات أخرى لا توفرها الاختبارات مثل الوثائق والسجلات المدرسية ونتائج الامتحانات السابقة ، وهذه البيانات وغيرها لابد من توافرها حتى يكون الحكم على مدى تحقيق الأهداف التربوية سليما .

وعند اعطاء الاختبارات وتطبيقها لابد من انتقاء العينات موضع البحث انتقاء سليما ، كما لابد أن تراعى الشروط الخاصة باعطاء الاختبارات وبهذا تبدأ عملية الحصول على المعلومات والبيانات عن الجوانب التربوية المختلفة موضع التقويم .

وبعد الحصول لى هذه البيانات من مصادرها الصحيحة لابد من تنظيمها بطريقة تجعل الاستعانة بها والاستفادة منها مسألة يسيرة تساعد فى اتخاذ القرارات الملائمة له .

وقد يحتاج الأمر الى تنظيم برنامج للتدريب الدورى لمن يقومون باعطاء الاختبارات على الأساليب الجديدة وعلى الاختبارات الجديدة التى يمكن الاستفادة منها فى البرنامج التقويمى ، ويدخل فى هذا تدريب المدرسين الذين يبدون اهتماما خاصا على استعمال وتطبيق بعض الاختبارات والطرق المستحدثة .

٣ - اصدار الأحكام : فى ضوء المقارنة بين البيانات التى يحصل عليها المسئول من التقويم باستخدام الأدوات الملائمة والأهداف التربوية المحددة يتم اصدار الأحكام ، وبالطبع فان هذه الأحكام قد لا تكون مهمة خبير التقويم منفردا ، وانما قد يشاركه فيها عدد من المسئولين فى المنظومة التعليمية ، ومنهم المتعلمون أنفسهم . وتزداد هذه المقارنة يسرا اذا كان الهدف قد صيغ فى عبارات اجرائية ، لأن هذه الصيغة تتضمن تحديدا لمستوى

الجودة أو الاتقان أو التمكن ، بالإضافة الى أنها تسهل لنا اختيار الأدوات التى ترتبط ارتباطا مباشرا بالهدف .

٤ - اجراءات التطوير : تعد خطة التطوير والتنمية والتحسين والعلاج فى ضوء الأحكام التى يتم الوصول اليها ، وقد تشمل قرارات التطوير تقديم برنامج فى التربية التعويضية أو التعلم العلاجى أو ايداعا فى احدى المؤسسات أو التدخل بأصاليب الارشاد النفسى أو العلاج النفسى المختلفة .

خصائص البرنامج التقييمى :

هناك سمات وخصائص لابد أن تتوفر فى أى برنامج تقييمى تربوى حتى يحقق الأهداف المرجوة منه يؤدى وظيفته بنجاح (٦ : ٣٨ - ٤٦) .

وأهم هذه الخصائص :

١ - الشمول : يجب أن يراعى البرنامج التقييمى الشمول ، وبيان هذا أنه لا ينبغي أن ينحصر الاهتمام فى القياس على المعارف والحقائق والمفاهيم . بل يتسع ليشمل الاتجاهات والميول والتفكير الناقد والتوافق الشخصى والاجتماعى ، وأن تستعمل شتى الأساليب بتوازن ، والحق انه اذا كان اهتمام المشرف على عملية التعليم يهتم بالتلميذ اهتماما كليا شاملا ، فمن اللازم ان يكون اهتمام البرنامج التقييمى ايضا شاملا لكل جوانب التلميذ وأنشطته .

٢ - التنظيم : يجب أن تجمع نتائج الاختبارات وببياناتها وتنظيم وتحلل بحيث يصبح تفسيرها ممكنا وذا مغزى ، فنتائج البرنامج التقييمى سواء كانت كمية ، أى فى صورة درجات أو مذكرات صغيرة أو ملاحظات عابرة ، أو كيفية ، يجب أن تلخص فى اطار واضح من التقديرات وأن تتحول الى صورة احصائية أو رسوم بيانية أو تقارير لغوية . وأن توصف اتجاهات هذه النتائج باستخدام الطرق الملائمة من التحليل الاحصائى أو تحليل المحتوى ، وأن تقدم فى جميع الأحوال فى صورة قابلة للفهم والنقل والاتصال بحيث تعطى صورة عن الفرد يمكن مقارنتها بصورة سابقة له ، فيساعد هذا

على التعرف على اتجاهات نموه فى المجالات المختلفة ، وكذا مقارنتها بصور زملائه الآخرين .

٣ - التكامل : لابد أن تتكامل المعلومات فى استعمالها عند القيام بحكم معين ، وعندما نضع خطة عمل بناء على هذا الحكم ، فالمعلومات الخاصة بالنواحى الصحية والتوافق العاطفى والاجتماعى والاتجاهات وغيرها يجب أن تربط مثلا بنتائج اختبارات التحصيل أو اختبار تحصيل بعينه . وكل ما لدينا من بيانات ومعلومات عن الفرد لابد أن نعامله باعتباره وحده يوضح ويكمل بعضها بعضا .

٤ - الاستمرار : يتميز البرنامج التقييىى الناجح بالاستمرار ، أى أنه فى المؤسسة التربوية الحديثة يعتبر عملية مستمرة . فالملاحظات اليومية والتقديرات والاختبارات المستمرة هى التى تشكل العمليات التقييىية ، وعن طريقها يحاول المدرس أن يقوم نمو التلاميذ ويوجهه . فلا ينبغى أن تكون هناك فترة محددة للتقويم ثم يتوقف بعدها أو ينتهى تقويم التحصيل المدرسى مثلا فى يوم حساب معلوم أو فى أيام معدودة يوضح فيها الميزان . والحكم الذى ينطق به لا يفيد المتعلم بأن يذله على مواطن الضعف أو القوة . أو الصحة أو العلة فى أسلوب التعلم والتحصيل والمراجعة ، فلا يفيد من هذا الحكم متعلم أو معلم وإنما هو قول فصل . أما النهج القويم فى تقويم التحصيل خاصة ، فهو أن يتداخل مع عملية التعلم نفسها ، يسبقها ، ثم أثناءها خطوة خطوة حيث يجرى القياس والمقارنة ويكون حقا قولنا أن التقويم السليم ضرورى للتعلم السليم ، بل وضرورى لفاعلية المؤسسات التربوية على اختلاف أنواعها ومستوياتها ، ويتمثل ذلك خاصة فى الاهتمامات الحديثة بما يسمى التقويم التكوينى .

٥ - الوظيفية : لابد أن يعرف كل أعضاء المؤسسة التربوية التى يعمل فيها البرنامج التقييىى بما يمكن أن يقدمه لهم هذا البرنامج من خدمات . فيعمل مثلا على أن يعرف التلاميذ والمعلمون ، بل وأولياء الأمور ، أنه من الممكن اعطاء اختبارات خاصة غير تلك التى تطبق على كل التلاميذ ، أو يمكن اعطاء اختبارات اضافية اذا أراد التلميذ أن يتعرف على بعض جوانب شخصيته . وربما كان أخذ مثل هذه الاختبارات بابا يدخل منه التلميذ الى

مناقشة بعض مشكلاته فيما يتصل بعلاقته بزملائه أو بمدرسيه أو بعمله المدرسي أو حياته الأسرية .

البرنامج التقويى اذن دعامة رئيسية من دعائم العملية التربوية ، ويعتمد نجاحها الى حد كبير على نجاحه ، لهذا كان من الواجب على المسئولين عن العملية التربوية ان يعطوا البرامج التقويمية التربوية كل جهد أو مال تستحقه ، وهو جهد ومال مبدول فى خير وجه ، وانه لمعط ثماره أضعافا مضاعفة .

دور الاختبارات فى البرنامج التقويى :

على الرغم من أن دور الاختبارات فى البرنامج التقويى يبدو لنا متواضعا ، حيث يحتل خطوة واحدة من خطوات هذا البرنامج ، الا انه فى واقع الأمر أكبر أثرا واشد خطرا من ذلك ، ونكاد نقول أن الاختبارات تكاد تكون القاسم المشترك بين معظم مكونات البرنامج التقويى ابتداء من الأهداف وانتهاء بالأساليب التطويرية والعلاجية ، وفيما يلى نعرض لبعض الأدوار التى تقوم بها الاختبارات فى مختلف جوانب هذا البرنامج ويشمل ذلك : الأهداف واتخاذ القرارات والاجراءات التطويرية وخاصة الارشاد النفسى .

أولا : دور الاختبارات فى صياغة الأهداف :

تساعد الاختبارات فى اختيار الأهداف التربوية وتوضيحها على نحو غير مباشر ، فهى تحفز المؤسسة التربوية على صياغة أهدافها والتعبير عنها بوضوح فى عبارات سلوكية أو اجرائية ، الا ان هذا التحديد السلوكى الواضح فى ذاته ليس ضمانا لاختيار أهداف تربوية صالحة . فعلمية اختيار الأهداف ليست من العمليات البسيطة التى تعتمد على مجرد التعرف أو الحكم المباشر ، وانما هى من الأمور التى تتطلب المفاضلة بين أنظمة قيمية ومعارية . ومع ذلك فهناك مجموعة من المعلومات التى يمكن للمدرسة أو المعلم أن يحصل عليها مما يهيئ فرصا أفضل لاختيار أكثر حكمة للأهداف . ومجموعة المعلومات هذه كما يصنفها رالف تايلر (٣٢) هى :

١ - مجموعة من المعلومات تتعلق بالافراد انفسهم قدراتهم ومعلوماتهم (التقويم النفسى)

ومماراتهم وميولهم واتجاهاتهم وحاجاتهم ، وبالطبع ترتبط هذه الخصائص بالمرحلة النمائية التي يمر بها هؤلاء الأفراد ومطالبها في نواحي النمو الجسمي والفسولوجي والعقلي والانفعالي والاجتماعي والمهني . ومراعاة هذه الخصائص في صياغتنا للأهداف يجعل عملنا التربوي معينا للفرد على تشخيص صعوباته وعلاج مشكلاته واشباع حاجاته ومواجهة مطالبه .

٢ - مجموعة من المعلومات تتعلق بمطالب المجتمع . ويتطلب ذلك معرفة جوانب القوة والضعف التي يؤثر بها المجتمع الحديث على عمليات التربية والتدريب عموما بحيث يكتسب الفرد من العملية التربوية المعلومات والمهارات والتدريبات والاتجاهات والميول التي لها أهمية في المجتمع والتي تساعد على التكيف مع الثقافة وتحسينها .

٣ - مقترحات المتخصصين في مختلف ميادين المعرفة والتي يرون بها كيف يمكن لمياديزهم أن تساهم في عمليات تعليم واعداد وتدريب الأفراد .

هذه الأنواع الثلاثة من المعلومات قد توحى بكثير من الأهداف التي قد تتعارض فيما بينها وبالتالي يصبح من الضروري القيام بالاختيار من بينها وهذا الاختيار يعتمد أساسا على فلسفة المدرسة أو المؤسسة التربوية أو برنامج التدريب الذي نقوم به ، والتي يمكن بها تحديد الأهداف ذات القيمة الكبرى في ضوء مفاهيم « المواطن » و « المجتمع » . كما يعتمد هذا الاختيار على نتائج الدراسات والبحوث في سيكولوجية التعلم . وفي هذه الأحوال جميعا تلعب الاختبارات دورا هاما في تزويدنا بالمعلومات حول الأفراد مما يوحي بأهداف معينة . وبعبارة أخرى يمكن القول بأن الاختبارات تلعب دورا مزدوجا في هذه الخطوة الأولى من عملية بناء خطط التعليم والتدريب : تستشير الهيئات التعليمية لاختيار وتحديد أهدافهم بوضوح حتى يمكن استخدام الاختبارات في تقويمها ، كما أنها ذات فائدة كبيرة في الحصول على المعلومات التي نستخدمها عندما نتخذ قرارا باختيار أهداف معينة دون غيرها .

ويتوافر للاخصائيين في التربية في الوقت الحاضر مصادر هامة للأهداف قام باعدادها المهتمون بالتقويم والقياس النفسي والتربوي ، سواء

فى المجال المعرفى (٢٠) أو الوجدانى (٢٦) أو الحركى ، ويمكن الرجوع الى مصدر آخر للحصول على مزيد من التفاصيل (١١) .

ثانيا - دور الاختبارات فى اتخاذ القرارات :

كل من يتعامل مع الناس يقوم باستمرار باتخاذ قرارات فيما يختص بسلوك هؤلاء الناس . وعملية اتخاذ القرارات بشأن الآخرين موجودة فى تعاملنا الاجتماعى بعضنا مع بعض ، حتى فى مواقف الحوار بين شخصين ، حيث يتخذ أحد المتحدثين قرارا بتغيير الموضوع لأن الشخص الآخر بدا عليه من العلامات ما يوحى اليه بأن موضوع الحديث مؤلم له أو غير مرغوب فيه على أقل تقدير ، وكذلك بالنسبة للمواقف الاجتماعية الأخرى الأكثر تعقدا . كما أن عملية اتخاذ القرارات تحتل مكانا هاما فى الشؤون الادارية وسياسة الأفراد حيث تقتضى طبيعة العمل اتخاذ قرارات فيمن يقبلون وفى توزيعهم على تخصصات مختلفة وترقيتهم بل والاستغناء عنهم . وكذلك نجد لاتخاذ القرارات أهميتها التى نحس بها جميعا فى مجال العملية التربوية ، فالمدرس يتخذ قرارا خاصا بمدى استعداد تلاميذه للانتقال الى موضوع دراسى جديد ، والمدرسة تتخذ قرارات فيما يختص بتوزيع التلاميذ ، ونقلهم بل وفى معاملة من يصدر عنهم مشكلات سلوكية ملحوظة ، أو من يسعون الى حل مشكلاتهم لدى المتخصصين فى المدرسة مثل المشرف الاجتماعى ، أو الموجه التربوى ، أو رائد الفصل أو ناظر المدرسة أو أى مدرس يثق فيه التلميذ صاحب المشكلة .

وكلما كان لدى الشخص الذى يتخذ القرارات معلومات كافية وصحيحة الى حد بعيد عن الفرد أو الأفراد الذين يتخذ قرارا بشأنهم كانت الفرصة أمامه أرحب لأن يحصل على أطيب النتائج وأنسبها ، فتوفر المعلومات قبل اتخاذ القرار يساعد متخذ القرار على أن يصدر أحكام قرار ممكن .

والاختبارات ، على اختلاف أنواعها نفسية أو تربوية أو طبية هى وسائل لاعطاء معلومات لا نستطيع أن نحصل عليها بغيرها ، وهذه المعلومات تساعدنا فى اتخاذ القرارات .

وإى قرار يتضمن تنبؤا ، ذلك لأن الاختبارات كلها تدلنا على بعض

الاختلافات في أداء أفراد معينين لحظة اجراءات الاختبارات عليهم ، ولكن هذه الحقيقة تكون عديمة القيمة اذا لم تستطع ان تتنبأ ان هؤلاء الأفراد يختلفون في أداء آخر ، أو في نفس الأداء ، في وقت آخر . وللاختبارات التي تتنبأ أهمية في النواحي التعليمية ، وفي العلاج النفسي وفي البحث العلمي .

وتستعمل الاختبارات لاتخاذ قرارات من انواع كثيرة ، منها : انتقاء الأفراد وتصنيفهم ، وتقويم طرق المعاملة المختلفة (طرق تدريس مثلا) ، وقبول أو رفض الفروض العلمية . وسوف نتكلم عن كل من هذه الأنواع من القرارات فيما يلي :

١ - الانتقاء والتصنيف : الانتقاء عبارة عن قبول أو رفض ، فعندما نعطي اختبارات متنوعة لمجموعة من الشبان المتطوعين للخدمة العسكرية أو المجندين ، ثم نقرر قبول بعضهم ورفض البعض الآخر حسب نتائج هذه الاختبارات المتنوعة ومستعنيين بكل ما يمكننا الحصول عليه من معلومات خاصة بهؤلاء الشبان ، يكون عملنا هذا انتقاء أما التصنيف فيقوم على تحديد المعاملة التي يتلقاها المتلقى ، فإذا نظرنا فيما سيتم بشأن الشبان المقبولين للعمل في الخدمة العسكرية فنجد أن منهم من سيوزع على فرع الطيران ، أو الميكانيكا أو المشاة أو غيرها ويتلقى تدريباً أو معاملة مناسبة لهذا التصنيف ، وكذلك نجد أن الشخص المصاب باضطراب انفعالي (أو مرض نفسي) يتحدد نوع العلاج الذي يتلقاه بناء على التشخيص المبني على الاختبارات المختلفة ، وهذا التشخيص عبارة عن تصنيف نوع الاضطراب أو المرض ، ويساعد التصنيف في التركيز على نوع المعاملة العلاجية المناسبة لهذا الشخص والتو تؤدي الى أحسن النتائج معه .

ومن التصنيف أيضاً تعيين الأفراد على « مستويات » للعمل وليس فقد لأنواع العمل . فإذا كان من نتائج اختبارات اعطيت للتلاميذ في اللغة ان قمنا بتوزيعهم على مستويات تتراوح بين الضعيف والمتوسط والحصن ، ليتلقوا تدريباً لغوياً مناسباً ، كل حسب مستواه ، كان هذا تعييناً للأفراد على مستويات للعمل .

ومن الضروري أن نميز بين اختبارات الانتقاء واختبارات التصنيف .
فإذا أعطينا اختبارات كانت نتيجتها استبعاد من عندهم اضطراب انفعالي،
مثلا كانت هذه اختبارات انتقاء ، لأن القرارات التي تبني على المعلومات التي
نحصل عليها منها قرارات قبول أو رفض فحسب ، لا تدلنا على أكثر من مدى
الصلاحية أو عدم الصلاحية للقبول في فئة أو مؤسسة ما . أما الاختبارات
التي تعطى معلومات تساعد على تحديد أو التعريف بطريقة معاملتهم أو تدريبهم
أو علاجهم فهي اختبارات تصنيف .

ومن المهم أن نشير إلى ضرورة الوصف الفردي المفصل في عملية
التصنيف ونعني بهذا ضرورة الحصول على معلومات مفصلة قائمة على
اختبارات متعددة ومن مصادر مختلفة عن الفرد ، من ناحية خصائصه
الشخصية المميزة ، واتجاهاته وميوله . واجتماعيته أو طبيعة علاقاته
الاجتماعية ، إلى جانب نتائج الاختبارات المحددة التي تهدف إلى تصنيف
معين ، فمن اللازم أن تكون تحت أيدينا مثل هذه المعلومات عن الطالب الذي
يرغب الدخول في كلية الهندسة أو الطب إلى جانب نتائج اختبارات
الاستعداد .

(٢) تقويم المعاملات : إذا كانت للاختبارات أهميتها في اتخاذ
قرارات كما بينا منذ قليل ، فإن لها أهميتها أيضا في تقويم المعاملات ، أي كان
نوع هذه المعاملات تدريبا أو تدريسا أو علاجيا . ولعله يكون واضحا أن ما
نعنيه بالمعاملة هنا هو الطريقة التي يعامل بها فرد أو أفراد أي كان مجالها
أو نوعها . فالاختبارات التحصيلية مثلا ، أو اختبارات التحصيل المدرسي ،
يمكن استخدامها لمقارنة طريقتين من طرق تدريس مادة الرياضة في المدرسة
الاعدادية للتعرف على أفضل الطرق في تدريس هذه المادة . ويمكن الحكم
على سياسات الاشراف والأفراد ومقارنة فاعليتها عن طريق قياس الاتجاهات
وللروح المعنوية بين الأفراد .

(٣) التحقق من الفروض العلمية : كثيرا ما تستعمل الاختبارات
النفسية بأنواعها المختلفة في التحقق من الفروض العلمية التي نضعها في
مجالات علم النفس المختلفة . والفرض العلمي بصفة عامة عبارة عن توقع
معين أو تحديده لعلاقة وظيفية بين متغيرات ، ونستخدم الاختبارات للتحقق عن

مدى صحة هذا الفرض أو خطئه . فإذا أقمنا فرضا علميا يذهب الى أن الأطفال الذين لا ينشأون في وسط عائلي طبيعي بين والديهم وأخوتهم يتميزون بدرجة ما من الاضطراب الانفعالي بعكس الأطفال الذين ينشأون في وسط عائلي طبيعي ، فإن من الممكن التحقق من صحة هذا الفرض باعطاء مجموعتين من الأطفال ، مجموعة نشأت في وسط عائلي طبيعي ومجموعة أخرى لم تنشأ في هذا الوسط ، اختبارات شخصية تدلنا على درجة اتزانهم الانفعالي ونقارن نتائج المجموعتين بعضها ببعض الآخر . وكذلك إذا كان لدينا فرض علمي يقول أن الأطفال يتعلمون بدرجة أحسن إذا تعاونوا في مواقف التعلم فإننا نستطيع التحقق من هذا باعطاء اختبارات تحصيلية متكافئة لمجموعتين من التلاميذ يتشابهون في كافة الخصائص ولكن مجموعة منهما تتعلم بالطريقة التعاونية بينما تتعلم الأخرى بطريقة مالوفة . ومن مقارنة نتائج الاختبارات التحصيلية للمجموعتين يمكن أن نختبر مدى صحة فرضنا . وهكذا في سائر المجالات النفسية والتربوية .

(ثالثا) دور الاختبارات في الارشاد النفسى :

الارشاد النفسى Counseling هو العملية التى يمكن بها تنظيم أو عرض المعلومات الخاصة بالفرد على نحو يساعده فى الوصول الى حلول فعالة لمشكلات التكيف التى يعانى منها والتى تكون عادة فى المستوى السلوكى العادى أو السوى . وبهذا المعنى فإن الارشاد النفسى جانب متخصص من جوانب الخدمات الفردية ، أو برامج التوجيه بوجه عام . أما إذا اتجهنا نحو مستوى الانحراف السلوكى الشديد فإننا ندخل فى باب علم النفس الطبى أو الطب النفسى والذى يتناول الأمراض النفسية والعقلية (كالعصاب والذهان) شديدة الخطورة سواء كانت ترجع لأسباب عضوية أو غير عضوية . وإذا اتجهنا نحو المستوى السلوكى الأكثر سوءا ندخل ميدان الخدمات البروتينية الخاصة بالتعريف المهنى واعطاء المعلومات عن المهن أو الفرص التعليمية . ولكن المشكلات الأبعد مدى من مجرد اعطاء المعلومات والأقل من المرض النفسى أو العقلى أو سوء التكيف الاجتماعى (انحراف الأحداث أو الاجرام) قد تكون مشكلات ملحة ومتنوعة وغالبا ما تكون على درجة من الصعوبة والتعقيد ، والأغلب تكون ذات أثر طويل المدى . ومن هذه المشكلات ما يتعلق بالتخطيط للمهنة والتخطيط لنوع التعليم اللازم للفرد ، ومشكلات التكيف

الاجتماعى والانفعالى التى يواجهها المراهقون سواء فى بداية مرحلة المراهقة أو نهايتها ، والمشكلات الخاصة بصعوبات التعلم والتى تتطلب تعليما علاجيا من نوع معين .

فالمرشد النفسى الذى يتناول هذه المشكلات وأمثالها بالبحث والدراسة والتشخيص والعلاج هو واحد من مجموعة كبيرة من المتخصصين فى الخدمات الفردية منهم الأطباء والأطباء النفسىون والاختصاصيون الاجتماعيون والمتخصصون فى عيادات الأطفال . والارشاد النفسى مثله مثل تلك التخصصات لا ينشط فى فراغ ولكن تحدده البيئة وظروف المجتمع التى تؤثر دون شك فى وسائل المرشد النفسى وأنشطته وأهدافه . فالنظام التعليمى عموما يتضمن أنماط من المناهج الخاصة التى تحددها ضرورات اجتماعية أحيانا أخرى . وبمثل الضرورة القانونية فى أن الطبيب لا يمكن أن يمارس هذه المهنة الا بعد التخرج فى كلية الطب ، على حين أن اصرار أصحاب الأعمال على التخرج فى أقسام الصحافة وإدارة الأعمال مثلا يمثل الضرورة الاجتماعية عندنا . كما أن الأسر المصرية حين ترسل أبناءها الى المدرسة إنما تعتبر عملية التعليم فى ذاتها صورة متخصصة للتأمين ضد البطالة . وبالتالي فإن الآباء يلحون على أبنائهم أن يبذلوا قصارى جهدهم لمواصلة تعليمهم .

أما النظام التعليمى المصرى نفسه فإنه يؤكد أساسا الدرجات المدرسية ومجموعها ونظام القبول حسب مكاتب التنسيق التى أصبحت من الثوابت الرئيسية فى نظامنا التربوى ، والتى امتدت من القبول بالمدارس الاعدادية حتى القبول بالكليات والمعاهد العليا .

وإذا كان لموظيفة الارشاد النفسى أن تحقق أغراضها فى مجتمعنا فلا بد أن تعمل خلال المحددات التى تضعها هذه الاتجاهات والمطالب والتقاليد الاجتماعية . وسواء اتفقنا أو اختلفنا معها فإنها تمثل الإطار المرجعى القائم الذى يعمل المرشد النفسى فى حدوده ، تماما كما يكون على الأطباء أن يتناولوا مشكلات الصحة الصناعية لصناعة معينة فى منطقة معينة . هذا من حيث الشكل الذى يحدد خدمات الارشاد النفسى . أما من حيث المضمون فإن هذه الخدمات وما يتعلق بها من تقويم وتياس تتأثر بأهداف المؤسسة (تربوية كانت

أم مهنية) وأغراضها . فإذا كان الهدف الرئيسى هو الانتقاء أو النصيح أو التزكية أو الاستشارة ، فإن معنى ذلك أن إجراءات الارشاد النفسى وعمليات التقويم والقياس المرتبطة بها تختلف عما لو كان الهدف الأكبر أكثر عمومية وأكثر اهتماما بنمو الأفراد وتحسين مستوياتهم .

والمرشد النفسى هو فى العادة أخصائى كلىنىكى يتعامل مع فرد واحد فى المرة الواحدة ، ولذلك فهو عادة ما ينزع الى أن يكون أكثر اهتماما بجوانب السلوك الفردى ، ملتزما بمعايير متعددة الأبعاد لتكيف الفرد . وهو لا يهتم بالنجاح كما يقاس بالتحصيل المباشر فحسب ولكن بالنجاح كما يقاس بالرضا الفردى والبقاء والاستمرار فى المدرسة أو المهنة ، كما أنه لا يحاول أن يتنبأ بنتائج برنامج معين فى القراءة العلاجية مثلا فحسب ولكن بالتكيف الاجتماعى للفرد كما يتأثر بالخبرات الاجتماعية والنمو . ومن واجبه ألا يركز على الصراع العاطفى المباشر فحسب ، ولكن على آثار هذا الصراع فى النمو العاطفى والانفعالى بوجه عام .

ولكن ما هو الدور الذى يقوم به التقويم والقياس فى خدمات الارشاد النفسى ؟

يمكن القول بأن علاقة التقويم بالارشاد النفسى هى علاقة الوسيلة بالغاية ، أى المساعدة على التعرف على نواحي القوة والضعف سواء كانت داخل الفرد أو بين الأفراد . بعبارة أخرى يفيد التقويم والقياس فى جمع المعلومات عن الأفراد الذين يعانون من المشكلات الفردية الخاصة بالتكيف والتوجيه والندى . كذلك مشكلات الجماعات من حيث التحصيل والانتقاء أو التوجيه المهنى والتربوى والتصنيف والتقسيم سواء فى الميادين التربوية أو المهنية أو العسكرية ، أى فى اتخاذ قرار ما . ولا يقتصر دور التقويم على مجرد جمع المعلومات وإنما يمتد الى زيادة استبصار المرشد النفسى وفهمه لسلوك الأفراد ولسلوكه هو أيضا حتى يستطيع الوصول الى التشخيص المناسب . كما يسمح التقويم والقياس ، وخاصة إذا توفرت فيه شروط معينة سنناقشها فى هذا الكتاب ، بالوصول الى تنبؤ أبقى مما لو استخدمنا أية وسائل لا تتوفر فيها تلك الشروط .

والواقع أن التشخيص والتنبؤ المعتمدين على الوقائع لها أهمية خاصة في حياتنا الاجتماعية عموماً . فمثلاً لا يمكن الاعتماد على أى برامج للانتقاء أو التوجه إلا إذا سمحت هذه البرامج بالتنبؤ الكمي الجيد ، وعلم الطب يعتمد أساساً على التشخيص من ناحية وعلى القيمة التنبؤية للمقاييس الفسيولوجية من ناحية أخرى . كما أن كلا من الطب النفسي والخدمة الاجتماعية قد أفاد كثيراً من المعلومات الكمية التي نجمتها عن سلوك الأفراد .

وهذا المنطق يصدق بالطبع على خدمات الارشاد النفسي . وهناك اتجاه يرى استخدام الاختبارات بأنواعها مسألة أساسية في العلاج والارشاد النفسي ، ذلك لأنها أدوات يتحدد في ضوئها تشخيص نوع الصعوبة الذي يعاني منها العميل ، ومنها يستطيع المعالج أو المرشد النفسي أن يكون فكرة أقرب إلى الدقة عن هذا العميل . كما أن الاختبارات - وخاصة اختبارات الشخصية - تساعد في التعرف على مدى التقدم الذي يحرزه الشخص في أثناء عملية الارشاد النفسي أو العلاج النفسي .

ويمكن أن نلخص الأسباب التي من أجلها يؤكد أصحاب هذا الاتجاه أهمية الاختبارات في عملية الارشاد النفسي فيما ذكره بوردين (٢١ : ٢٦٢ - ٢٧٩) وتايلور (٣٣) كما يلي :

١ - تعد الاختبارات المرشد النفسي بمعلومات معينة يحتاج إليها في عمله مع الشخص وتساعد في مهمته معه .

٢ - تعد الشخص أو العميل بمعلومات معينة يحتاج إليها وخاصة في اتخاذ قرارات معينة .

٣ - قد تستثير الاختبارات العميل لاستكشاف ذاته .

٤ - أن العميل قد يكتسب استبصاراً وفيها لنفسه وهو يقوم بأداء الاختبار .

الا ان هناك اتجاها آخر ينظر الى استعمال الاختبارات فى العلاج والارشاد النفسى نظرة مغايرة ، اذ يتردد أصحاب هذا الاتجاه فى الاعتماد كلية على منهج التقويم والقياس . وهذا التردد يعتمد على حجتين أساسيتين: أولاهما أن تأكيد المنطق الكمي والطبيعية العقلانية لعملية الوصول الى استبصار بموقف العميل قد يؤدي الى تضليلنا ، لأن ما نعتبره معلومات موضوعية محايدة هو فى حقيقة الأمر مرتبط بالكيان الانفعالى والعاطفى للعميل ، وثانيهما أن معرفتنا الراهنة بالشخصية وأدوات تقديرها جميعا لازالت فى مراحلها الأولية المبكرة .

وهذا الاتجاه يمثل فى الوقت الحاضر كارل روجز Carl Rogers (٢٨ ، ٢٩) وتلاميذه حيث يركزون الاهتمام على الجوانب العلاجية فى الارشاد النفسى ويتبعون طريقة خاصة من طرق العلاج تتميز بعدم التأكيد على الاختبارات لأنهم لا يقيمون وزنا كبيرا لعملية التشخيص والتصنيف فى العلاج . وجوهر طريقتهم هو أن الفرد مركز التقويم . أى أن التقويم يأتى من داخل الفرد ، وأن نموه وشفاءه أو حل المشكلة التى يعانى منها انما يأتى من قوى داخلية تنبعث منه عندما يوفر له المعالج الجو المطمئن الخالى من أى نوع من أنواع التهديد للذات أو التقويم للفرد . هذا الجو الذى يتحقق عندما يدعم المعالج عند العميل باستمرار وبعشق أنه يتقبله ويثق به ويهتم به باعتباره انسانا فى أزمة وجود أو أزمة حياة . ويساعده على التغلب عليها ليستمر نموه .

هذا بايجاز شديد هو الجوهر الذى يدور حوله العلاج المركز حول العميل ، وفى ضوءه نفهم الأسباب التى جعلت الاختبارات عند أصحاب هذا الاتجاه لا تحتل مكانا هاما فى العلاج أو الارشاد النفسى ، ويمكن أن تلخص هذه الأسباب فيما يلى :

(١) المجال الذى ينشط فيه الفرد هو على درجة كبيرة من الاتساع وأن مواقف الأداء الفعلى عديدة ومتغيرة بحيث لا يمكن تمثيلها فى أى موقف اختبارى (الذى هو فى ذاته بحكم تعريفه ليس الا عينة من السلوك) .

(٢) أن تأثير المتغيرات والعوامل التى يصعب أو يستحيل قياسها فى

الوقت الخاضر قد تكون على درجة كبيرة من القوة والأهمية ، مثل عوامل الحوافز والدوافع • الى حد أن الصفات السلوكية التي يمكن قياسها - ان قورنت بما لا يمكن قياسه - قد تكون اقل اهمية بالنسبة للنشاط الواقعى .
فى للحياة (١٧) •

(٣) الاختبارات ليست الا عملية مقارنة بين العميل وآخرين فى جانب من الجوانب ، بينما المشكلة التي يعانى منها العميل هى فى حقيقتها مشكلة ذاتية نابعة من ادراكه الخاص بالمواقف والأشخاص ، ومن ادراكه لنفسه ، ومن المعانى التي يخلعها على هذه المدركات جميعا ، ولا يمكن ان نفهم العميل ومشكلته الا فى ضوء هذه الذاتية •

(٤) توحى الاختبارات الى العميل أن دور المعالج أو المرشد النفسى هو أن يقوم بعمل شئ له (أى للعميل) يحل مشكلاته وينهى معاناته ، ومن ثم يتخذ هذا العميل دورا سلويا ينتظر فيه ما سيعمل له بدلا من ان يحس مسئوليته فى علاج نفسه متعاوننا مع المعالج أو المرشد النفسى ، الذى لا يزيد دوره عن كونه - كما عبر روجرز - الأنا الأخرى للعميل • هذا الاتجاه السالب الذى تخلقه الاختبارات لدى العميل يجعله يتوقع « وصفة » لمشكلاته، ويصبح من أشق الأمور على المعالج أو المرشد النفسى نقل مسئولية الموقف واتخاذ القرارات الى العميل نفسه •

(٥) تحمل الاختبارات أيضا الى العميل شيئا من البعد بينه وبين المعالج أو تسبغ على العلاقة بينهما صفة الرسمية أو توحى بالعلو من جانب المعالج على العميل • وهذا النوع من العلاقة اذا جاز وجوده فى مجال الطب والأمراض الجسمية فانه غير مناسب فى علاقة المساعدة فى مجال الارشاد والعلاج النفسى •

(٦) قد يستعمل العميل الاختبارات عند أخذها وفى تفسيرها والتحدث حولها ستارا أو دفاعا أو هروبا من المشكلات الحقيقية التي ينبغى ان يواجهها مما يعطل سير عملية العلاج أو الارشاد النفسى •

بعض الوظائف التربوية للتقويم النفسى :

يؤدى التقويم النفسى وظائف عديدة فى مختلف مجالات الحياة كالصناعة والادارة والتربية والعلاج النفسى وغيرها ، على نحو يدل على أهميته وقيمه للباحث النفسى فى هذه المجالات ونعرض فيما يلى بعض هذه الوظائف فى مجال بذاته وهو مجال التربية ، ويمكن للمهتمين بالمجالات الأخرى أن يجدوا بسهولة ويسر وظائف التقويم النفسى واضحة لهم .

(أولا) دور التقويم فى زيادة الدافعية للتعلم :

يقوم التقويم بالوظائف الثلاث الرئيسية للدوافع فى التعلم وهى :

(١) وظيفة التنشيط energizing أى زيادة المستوى العام للنشاط والجهد المبذول ، ويمكن التذليل على ما تحدثه الامتحانات السنوية والعامية من ازدياد النشاط والجهد فى فترات العمل الشدید التى يقضيها طلاب المدارس الثانوية والجامعات قبيل الامتحانات . وكثيرا ما تكون هذه الامتحانات فى معظم المدارس خاتمة المطاف ، ويكون للنجاح والفشل فيها أهمية كبيرة . كما تحدد هذه الامتحانات عموما متى يتعلم التلاميذ وماذا يتعلمون وكيف يتعلمونه ، ولذلك فان هذه الامتحانات اذا لم تقم بقياس أهداف المنهج أو المقرر قياسا حقيقيا وصادقا فاننا نشك كثيرا فى قيمة الدافعية التى تستثيرها .

ويمكن القول من الوجهة المنطقية البحتة انه ما دام التقويم يزيد الجهد المبذول فان تكرار اجراءات التقويم يؤدى بالتالى الى زيادة الجهد الكلى للتلاميذ ، ولكن من المؤكد ان التكرار المناسب للتقويم يعتمد على طبيعة الأدوات التى تستخدمها ، وطرق التدريس ، وما اذا كان التلاميذ يصححون اختباراتهم أو اختبارات غيرهم ، والمادة الدراسية ، وقدرة التلاميذ بوجه عام . ولا شك ان اجراء التقويم بكثرة يقلل نسبيا من أهميته ، فالامتحانات اليومية مثلا ، قد لا يكون لها أى أثر منشط يزيد عن أثر الواجبات المدرسية اليومية . ويلخص كوك (٢٢) نتائج البحوث التى أجريت على أثر التقويم فى الدافعية فى أنه عندما تعطى كل أسبوع الاختبارات التحصيلية فى المواد

الدراسية في مستوى الجامعة ثم تناقش النتائج وتستخرج الأخطاء الفردية، وعندما تتم صياغة أسئلة الامتحان النهائي على صورة تشبه تلك الاختبارات فإن تحصيل الطلاب من ذوى المستوى المنخفض من القدرة يكون أعلى منه فى حالة التقويم الأقل تكرارا . أما الطلاب من ذوى المستوى المرتفع من القدرة فقد يتخلفون ما لم يكن فى كل اختبار أسئلة على درجة مناسبة من الصعوبة تتحدى قدرتهم ، كما أنهم يفيدون أكثر من المادة الإضافية التى يمكن للمعلم أن يدرسها اذا كرس وقتا أقصر للتقويم والعمل العلاجى بوجه عام . ولكن هل يمكن أن تعزى آثار التقويم المتكرر فى تحسين تحصيل التلاميذ الضعاف الى وظيفة التنشيط هذه ؟ الواقع أن معظم الباحثين يرون أن من المحتمل أن هذه الآثار ترجع الى توجيه التعلم واختيار الاستجابات الصحيحة (وهما الوظيفتان اللتان سنتحدث عنهما فيما بعد) أكثر من زيادة الطاقة والنشاط والجهد التى تستثيرها الامتحانات .

وقد تناول الباحثون بالدراسة أيضا اثر معرفة نتائج الامتحانات (والتقويم عموما) فى زيادة الدافعية للتعلم وأكدت نتائجهم أن معرفة النتائج أفادت التعلم بمقدار النجاح الذى تتضمنه النتائج ، وقد أفاد التلاميذ من ذوى المستويات المرتفعة أكثر من غيرهم أما أولئك الذين أحرزوا تقدما بطيئا فلم تستثرهم معرفة النتائج الا استثارة ضئيلة للغاية .

وقد درست هرلوك (فى ١١) اثر المدح العام أو اللوم العام فى التعلم . وأكدت نتائجها أن المدح أكثر استثارة للتعلم من الذم ، وهو أشد تأثيرا فى التلميذ الأقل نضجا والأقل قدرة ، أما الذم فكان أكثر فائدة بصفة عامة مع التلميذ الناضج والقادر ، وقد لوحظ أن كلا من المدح والذم أكثر استثارة للتعلم من موقف الحياد أو التجاهل التام .

وخلاصة القول ، أن الأثر المنشط للتقويم يعتمد أساسا على درجة نجاح التلميذ ، فالذى يحصل على درجات عالية وبالقالى ينجح ويتقدم ويلقى المديح تكون استثارته أيسر من التلميذ الذى يحصل على درجات منخفضة وما تستتبع ذلك من فشل ورسوب وذم ولوم . وهذه النتيجة ذاتها تؤكد ضرورة وضع أهداف تربية معقولة تتضمن على الأقل نجاحا جزئيا وممكنا للجميع .

(ب) وظيفة التوجيه directive ، وبها يتم توجيه سلوك المتعلم ونشاطه الى المسالك المرغوب فيها . فاذا كانت نتائج التقويم يمكن بها الحكم على المدارس والمعلمين والتلاميذ فان طبيعة ادوات التقويم تحدد الى حد كبير نوع العملية التربوية . فاذا كانت الاختبارات تتضمن أحكاما كيفية هامة فانها تفيد كثيرا في تحديد الأهداف والطرق التربوية ، واذا كانت الاختبارات تعتمد على مضمون تقليدي بحيث تدور معظم أسئلتها حول الحقائق بغرض تحديد مقدار ما حفظه التلاميذ من مادة الكتاب المدرسي ، فان آثار هذا النوع من الاختبارات تختلف عن الاختبارات التي تقيس مهارات الدراسة وقدرات حل المشكلة ، لأن الأخيرة تؤكد المفاهيم وتطبيقاتها في المواقف الجديدة وتوضح الأهداف الهامة وتركز الانتباه على الهدف الأقصى لعملية التربية وهو التعلم الدائم . ويمكن تلخيص الفرق بين نوعي الاختبارات في أن النوع الأول يجمد المنهج المدرسي والنوع الثاني يذيبه .

ولا شك في أن التقويم يؤثر في اجراءات الدراسة وأهدافها كما سبق أن أوضحنا . ففي اعداد امتحانات من نوع المقال يكون على التلاميذ تحديد المادة وتنظيمها في وحدات كبيرة وتأكيد العلاقات والاتجاهات والاستجابات الشخصية الخاصة بالتلميذ الفرد ، أما في اعداد الاختبارات الموضوعية التي تعتمد على الحقائق فانها تتناول عينة من السلوك اشمل من اختبارات المقال ولكنها تتطلب من التلاميذ تأكيد التفاصيل والاسماء والتواريخ ونتائج التجارب . وقد لوحظ أيضا بعض الاختلاف في طريقة اعداد التلاميذ لأنفسهم عند الاجابة على الأنواع المختلفة من الاختبارات الموضوعية ، فمثلا عند الاستعداد لاختبارات التكملة فان الطالب يحاول اتقان أسلوب الكلمة في مقابل الكلمة ، أما في حالة الاستعداد لاختبارات الصواب والخطأ او الاختيار من متعدد بوجه عام فان التلميذ يهتم بالتعاريف والحقائق التفصيلية . ولا شك ان هذا من العوامل التي تحدث فروقا فيما نتعلمه وفي مقدار ما نحفظه لفترة من الزمن وفي القدرة على اداء أنماط مختلفة من الاختبارات ، وكلها ترتبط تجريبيا بنوع الاختبار الذي يستعد الطالب للاجابة عنه .

(ج) وظيفة الانتقاء selective ، وتعنى هذه الوظيفة تحديد الاستجابات التي سوف يتم تثبيتها وبقاؤها عند التلاميذ والاستجابات التي

سوف تحذف ، وترتبط هذه الوظيفة بالجانب التشخيصى للتقويم ، الا ان الفرق بين الانتقاء والتشخيص هو ان الاختبار التشخيصى يركز الاهتمام على تشخيص الصعوبات كما يقترحها المعلم او الاختصاصى النفسى المدرسى ، اما ما نتحدث عنه الآن فهو ان التلاميذ والمعلم والاختصاصى النفسى جميعا يمكنهم الحصول على مقدار كبير من المعرفة من اخطاء التلاميذ (واسبابها اذا كان هذا ممكنا) فى موقف الاختبار . وينتج أكبر مقدار من التعلم من موقف الاختبار ذاته حينما يسمح للتلاميذ بتصحيح اوراق اجاباتهم ثم يلى ذلك مناقشة الأخطاء ثم العمل العلاجى . وتؤكد البحوث التجريبية انه كلما كانت معرفة التلاميذ لجوانب الصواب والخطا فى استجاباتهم واسبابها معرفة مباشرة فى موقف الاختبار فان ذلك يؤدي الى تثبيت الاستجابات الصحيحة او السلوك المرغوب فيه وحذف الأخطاء واستبعادها .

(ثانيا) دور التقويم فى تصميم المواقف التعليمية :

يعد مبدأ الفروق الفردية من المبادئ الأساسية فى علم النفس الحديث ، وهو يعنى من الوجهة التربوية التباين فى جماعات الاعمار الزمنية والفرق الدراسية فى الذكاء والقدرات والتحصيل والاتجاهات والميول وسمات الشخصية المختلفة . وأى رأى يعتبر محض الفرق الدراسية مرحلة لها خصائصها المحددة التى تنطبق على جميع الأفراد المنقسمين لها رأى خاطئ ويعطل جهود المؤسسة التربوية وسعيها لاشباع حاجات التلاميذ كأفراد . وأى افتراض حول وظيفة المعلم على أنه متخصص فى فرقة دراسية معينة انما يلغى امكانية نجاحه فى التعامل مع التلاميذ الذين قد يشذون عن المستوى الافتراضى لهذه الفرقة . وأى جهد من جانب السلطات التعليمية لجعل المقرر الدراسى جامدا ثابتا وعاما على جميع تلاميذ الفرقة الدراسية لا يضع فى الاعتبار أن هؤلاء التلاميذ ليسوا جميعا على نفس الدرجة من القدرة على أداء ما يطلب منهم بنجاح .

وقد أجريت عدة محاولات لحل مشكلة الفروق الفردية فى مواقف التعلم نذكر ما يلى :

١ - مراعاة القدرة العقلية العامة : من الافتراضات التى وجهت

الممارسات التربوية فى الماضى أن الجماعات المدرسية تكون أكثر تجانسا فى التحصيل إذا تجانست فى القدرة العامة ولذلك كانت تستخدم طرق تربوية مقننة مثل توحيد المقررات والامتحانات بحيث تتفق مع هذا المستوى العام ويتضمن هذا الافتراض مسألتين أولاها أن التباين داخل الفرد لا يعتد به ، وثانيهما أن جميع الخصائص السلوكية التى تهتم بها المدرسة ترتبط فيما بينها ارتباطا عاليا وتتنظم حول عامل يحدد المستوى العام لكفاءة الفرد . والواقع أن الأدلة العلمية الحديثة تدحض هذا الافتراض وتضميناته . وقد جاءت هذه الأدلة من مصادر عدة أشهرها البحوث الحديثة فى القدرات العقلية التى أكدت أن الاعتماد على القدرة العقلية العامة وحدها لا ينتج مجموعات متجانسة تجانسا كاملا (١٠) .

٢ - مراعاة مستويات التحصيل السابقة : من ناحية أخرى فإن المعلمين ورجال الإدارة التعليمية والمسؤولين التربويين يرجعون التفاوت الكبير فى مدى القدرة لدى المتعلمين الى عدم الدقة فى اجراءات نقلهم من المستويات التعليمية وهم فى هذا الصدد يطالبون بأن تكون للفرقة الدراسية الواحدة مستويات محددة يجب أن يصل اليها التلاميذ قبل انتقالهم الى فرقة أعلى ، وعلى الرغم من جودة هذه الفكرة إلا أنها صبت أحيانا فى القالب التقليدى حيث يحدد المقرر المدرسى ما يجب أن يتعلمه الجميع فى كل مادة فى كل فرقة دراسية وأى الكتب المدرسية تستخدم ، وأى صفحات أو فصول يجب أن تعطى فى وقت محدد من السنة ، وتكون محصلة الأمر أن يتعلم التلميذ المعارف المتضمنة فى هذه الكتب مثل : تهجى الكلمات ، المسائل النموذجية فى الحساب ، الأسماء التواريخ ، أسباب الحروب ونتائجها ، الدول وعواصمها ، والمدن وخصائصها الخ .

والسؤال الآن هو : هل توجد بالفعل مستويات للتحصيل لا يتباين فيها التلاميذ بحيث تؤدى الى تحقيق التجانس داخل الجماعات المدرسية ؟ للإجابة على هذا السؤال يمكن صياغة القاعدة العامة الآتية التى نستخلصها من نتائج البحوث العديدة التى أجريت فى هذا الميدان .

• إذا كانت الاستجابات التى نتعلمها بسيطة والأهداف التى ننشدها محددة بحيث يمكن لأكثر عدد من التلاميذ أن يصلوا اليها خلال فترة التعلم

فان تباين المجموعة يتناقض ، أما اذا كان العمل على درجة من التعقيد والتركيب والاهداف غير محددة بحيث لا يمكن أن يصل اليها الا التلاميذ الأقدر فان تباين المجموعة يتزايد » .

ومعنى ذلك أن الرأى القائل بأن التلاميذ يمكنهم أن يصبحوا أكثر تجانساً اذا استخدمنا أساليب التدريس التى تؤدى بهم الى مستوى واحد مقنن انما يفترض محدودية الاهداف وضيق نطاق التدريس فحتى يمكن أن تكون المهام فى متناول التلاميذ الأقل قدرة بلجأ المعلم الى أساليب التعلم الروتينية ، كما أن الكتب المدرسية تكون فى العادة من مستوى أدنى ، والمواد التى يتم اختيارها للمنهج تكون من النوع الذى يمكن أن يتقنه حوالى ٨٠٪ من التلاميذ ، ويتم تدريب التلاميذ على الأسئلة واجاباتها ، مثل أسئلة الامتحانات ، وعادة ما يكون معظم وقت المعلم مكرساً للتلميذ بطيء التعلم وفى هذه الحالة عندما يقاس تقدم التلميذ فى فترة زمنية معينة فى ضوء الاهداف المحدودة فان التلاميذ ذوى الدرجات المنخفضة فى البداية يحوزون أعلى تقدم ممكن ومعنى ذلك أن هناك علاقة سالبة بين الدرجات المبدئية وبين درجات الاستفادة أو الكسب فى نهاية التعلم . كما أن العلاقة بين الاستفادة فى هذه الحالة والذكاء علاقة منخفضة وكثيراً ما تكون سالبة . وهذه النتائج تؤكد أن مثل هذه الممارسة لا تستثير التلاميذ من ذوى القدرة العالية أو المتوسطة الا قليلاً ان لم تهملهم اهمالاً تاماً .

وقد تكون أخطر النتائج المترتبة على ذلك أن ما يتعلمه المتعلمون قليل القيمة ، ضعيف الأثر ، ضيقاً فى مدى الاحتفاظ به وتستخدم فيها أسئلة الحقائق والتفاصيل والمعلومات المحدودة التى تدور حول أذكر ، وصف ، وعرف ، ومن ، وماذا ، ومتى ، وأين . ولا يستخدم التلاميذ فى الاجابة عليها الا الذاكرة . وتكون آثار التعلم فى هذه الحالة مؤقتة وخاصة مع استخدام الذاكرة الضعفاء . لأن ما ينساه التلميذ فى هذه الحالة تكون ما بين ٤٠٪ خلال ثلاثة اشهر و ٨٠٪ خلال ثلاث سنوات ، ويسمى باكنجهام (فى ١٠) ذلك « اعظم فقد تربوى » . ومنحنيات النسيان لمعظم المواد الدراسية فى هذه الحالة تقترب كثيراً من منحنيات نسيان المواد عديمة المعنى .

أما اذا كان التعلم يتضمن عمليات حل المشكلة وغيرها من العمليات (التقويم النفسى)

العقلية العليا فان آثاره تكون أكثر دواما . أى أنه ما لم يعمل التعلم على تمايز وتكامل الاستجابات القديمة والجديدة فى نمط سلوكى منظم لا تكون لهذا التعلم قيمة أو لاهتمام . وهذه النتائج تهمنا كثيرا فى الوقت الحاضر مع انتشار مفهوم « التعلم للالتقان » وحتى لا يتحول هذا للشعار النبيل الى محض وصول الى مستوى متواضع من القدرة .

خلاصة القول أن حل مشكلة الفروق الفردية فى مواقف التعلم لا يكون بتحقيق تجانس التلاميذ فى القدرة العامة أو باستخدام الطرق المكننة وإنما بالاجابة على سؤال أهم هو كيف تم التعلم وكيف صممت مواقفه ؟ فإذا سادت التعلم العمليات العقلية العليا وكانت أهدافه من العمومية والاتساع والارتفاع أدى ذلك الى تحقيق نتائج أكثر استمرارا ودواما والى استثارة المتعلمين لبذل قصارى جهدهم فى التعلم فأننا بذلك نصل بالمتعلمين الى الحد الذى يظهر قدراته وتفرداه وعلى نحو فيه التباين الابداعى أكثر مما فيه من التجانس الجامد .

(ثالثا) دور التقويم فى تصنيف المتعلمين :

لجا بعض الباحثين فى حلهم لمشكلة الفروق الفردية الى افتراض أن التباين داخل الفرد والتباين بين الأفراد يدعوان الى ضرورة مقابلة الحاجات الفردية . ويعنى ذلك أن نقبل حقيقة وجود التباين الواسع فى السلوك الانسانى .

ولمواجهة هذا المطلب فى المدرسة الابتدائية مثلا اهتم الباحثون بخصائص نمو التلاميذ من ناحية ، وتكيفهم المناسب للثقافة التى يتأثرون بها من ناحية أخرى . وفى عملية التفاعل الاجتماعى العام خلال سنوات التكوين تكون السمات الأكثر أهمية فى تحديد عضوية الطفل ومكانته فى الجماعة هى للعمر الزمنى والنمو الجسمى العام ، وإذا أضفنا الى ذلك النمو الاجتماعى العام ، تكون هذه السمات الثلاث فى رأى المهتمين أساس التصنيف والتجميع التربوى . الا أن ما يجب ملاحظته أن العمر الزمنى لا يرتبط ارتباطا كاملا بالنمو الجسمى والاجتماعى للطفل . ولهذا تعطى الأولوية لمستوى النمو

لا العمر الزمنى وبخاصة النمو الجسمى والاجتماعى ، حتى يمكن للطفل أن يعيش ويعمل فى مجموعة ينتمى إليها تقبله ويقبلها .

ويضاف الى عامل النمو (وخاصة النمو الجسمى والاجتماعى) فى التصنيف باعتباره العامل الأولى والأساسى عامل آخر هو مستوى التحصيل النوعى فى المجالات المختلفة تبعاً لمقدرات التلاميذ فى مختلف المواد الدراسية . ويجب أن نتذكر دائماً أنه حتى لو تم التجميع والتصنيف فى المدرسة الابتدائية بهذه الطريقة فلا بد أن يتركز اهتمامنا على التلميذ الفرد . وبالطبع فإن ملاحظات المعلم لتلاميذه وحكمه عليهم يكون لها دائماً أهمية قصوى فى موقف التعلم .

دور الحاسب الألكترونى فى التقويم النفسى :

يتزايد فى الوقت الحاضر استخدام الحاسب الألكترونى (الكمبيوتر) فى المجالات المختلفة لعلم النفس سواء داخل العمل أو خارجه . ولعل أكثر هذه المجالات استخداماً للحاسب الألكترونى مجال التقويم النفسى ، ومن المتوقع أن يزداد هذا الاستخدام شيوعاً وانتشاراً فى المستقبل .

وتمتد الاستخدامات الراهنة للحاسب الألكترونى على نطاق عملية التقويم النفسى كلها ، إلا أن أكثر هذه الاستخدامات وضوحاً فى عملية تصحيح الاختبارات وتحليل نتائجها . وتتوافر لدى ناشرى الاختبارات فى الوقت الحاضر إمكانات تزويد الباحثين بنتائج الاختبارات على أى صورة يشاءون ابتداء من نتائج المؤسسات وحتى نتائج الأفراد ، وذلك بعد تصحيح أوراق الإجابة التى عادة ما تتخذ صورة معينة تصلح للمعالجة بواسطة الكمبيوتر . ويستخدم الحاسب الألكترونى أيضاً بسهولة ويسر فى عمليات تحليل أسئلة الاختبار وانتقاء الملائم منها كما يمكن المقارنة بين المعايير المحلية والمعايير القومية للاختبار ، ويستطيع الحاسب الألكترونى أيضاً أن يزود المدارس والمؤسسات المختلفة بنتائج الاختبار لتقوم بدورها فى التغذية الراجعة ، وقد يسر ذلك على المعلمين والاختصاصيين أن يستثمروا هذه النتائج فى تحسين برامج التعليم والتدريب .

ويتزايد فى السنوات الأخيرة استخدام الحاسب الالىكترونى فى عمليات اعداد الاختبار ، وخاصة الاختبارات التحصيلية ، واستخدامها داخل المؤسسات نفسها (كالمدارس) ومع ظهور الحاسبات الالىكترونية المصغرة تتوافر فى الوقت الحاضر اعداد منها فى معظم المدارس الأوربية والأمريكية، تفيد المعلمين فى سرعة تصحيح الاختبارات التى يقومون باعدادها ، ناهيك عن الحاسبات الالىكترونية الكبيرة التى تستخدم فى تصحيح وتحليل نتائج الاختبارات ابتداء من المستويات المحلية وحتى المستوى القومى العام .

وظهرت فى السنوات الأخيرة أيضا نظم أكثر تقدما تشمل ما يسمى « بنوك الأسئلة » أو « مخازن مفردات الاختبارات » ، والتى تقوم فيها الحاسبات الالىكترونية بالدور الحاسم ، وتفيد هذه «البنوك » أو « المخازن » فى انتقاء المفردات أو الأسئلة لاعداد اختبار بمواصفات معينة (مثل أهداف تعليمية معينة ، محتوى معين ، مستوى صعوبة معين ، الخ) . وعادة ما يشمل النظام الكامل لبناء الاختبارات عن طريق الحاسب الالىكترونى المكونات الآتية :

- ١ - بنك أسئلة .
- ٢ - نظام لتوليد الأسئلة .
- ٣ - بنك بخصائص الأسئلة .
- ٤ - نظام لانتقاء الأسئلة .

ومن المجالات الهامة التى امتد اليها استخدام الحاسب الالىكترونى فى التقويم النفسى مجال تطبيق الاختبارات ذاته . فمن الملاحظ أن عددا كبيرا من المدارس فى الولايات المتحدة واليابان وبعض الدول الأوربية المتقدمة تطبق الاختبارات المعتادة عن طريق الكومبيوتر وخاصة البرامج التعليمية التفريدية التى يديرها الحاسب الالىكترونى (أى حين يستخدم الكومبيوتر كمعلم) . ولعل مما هو أكثر أهمية استخدام الكومبيوتر فى تطبيق ما يسمى الاختبارات « التفصيل » ، tailored أو الاختبارات التكيفية التتابعية ، فالمرونة الفائقة التى عليها الحاسب الالىكترونى تسمح له بأن يطبق على كل مفحوص الأسئلة أو المفردات الأكثر ملاءمة لمستوى قدرته . وفى هذا الحالة يعتمد الكومبيوتر على استجابة المفحوص على سؤال معين ، ويتحدد فى ضوء هذه

الاستجابة نوع السؤال أو مجموعة الأسئلة التي تتلوه . ومثل هذه النوع من الاختبارات « التفصيل » توفر الوقت والجهد اللذين يبدلان في الإجابة على اختبارات شديدة السهولة أو الصعوبة بالنسبة للمفحوص .

ويوجد استخدام آخر للحاسب الإلكتروني يعد أكثر حداثة وهو تزويد الفاحصين ببعض التفسيرات المفيدة في الإرشاد النفسي ومن ذلك أن يقدم الكمبيوتر تقريراً لفظياً عن أداء المفحوص يصف المهن الأكثر ارتباطاً وملاءمة لاستعداداته وميوله وتطلعاته التربوية يسمى « تقرير التخطيط المهني » . وقد لوحظ أن الكمبيوتر يستطيع القيام بعمليات المقارنة بين أنماط درجات الاختبارات المختلفة للمفحوص ويقوم بالربط بينها وبين المتطلبات المهنية المختلفة على نحو أسرع وأكفاً ، من المرشد النفسي . ومن المتوقع أن يوظف في المستقبل القريب هذا الجانب من استخدامات الكمبيوتر في التربية وخاصة ما ينصل بتشخيص صعوبات التعلم المدرسي . وتوجد موارد لذلك في الوقت الحاضر بالنسبة لاختبارات المهارات الأساسية .

بعض المشكلات المعاصرة في ميدان التقويم النفسي :

يواجه ميدان التقويم النفسي في الوقت الحاضر عدداً من المشكلات تتصل في جوهرها باستخدام الاختبارات ، ويمكن أن نلخص هذه المشكلات فيما يلي :

١ - مستوى جودة الاختبارات :

شهد مطلع الستينيات بداية حملة منظمة ضد الاختبارات المقننة سرعان ما اشتدت وزاد سعيها في السبعينات ثم الثمانينات ، وقد وجهت سهام النقد على وجه الخصوص إلى الأسئلة الموضوعية من نوع الاختبار من متعدد التي تتألف منها معظم هذه الاختبارات ، واتهمت بأنها تعاقب المفحوص الأكثر ذكاءً وابتكارية . وقد دعم النقاد هذا الرأي بما يلاحظ من أن هذا المفحوص الذكي المبتكر يظهر في إجابته على أسئلة الاختبارات المقننة اختلافاً عن الإجابات التي تتحدد صحتها مقدماً والتي تتكون منها مفاتيح هذه الاختبارات ، وعادة ما يحكم على هذا النوع من الإجابات التي لا تتفق مع المفتاح بالخطأ بالرغم

من أننا لو حللناها بعمق وأناة نجد فيها قدرا كبيرا من الصواب . وهذا النقد بالطبع له ما يبرره وخاصة أنه وجدت في بعض الاختبارات المقتنة ذات الشهرة الكلاسيكية أسئلة من هذا النوع . ناهيك بالطبع عما يمكن أن تجده فيما يتوافر لدينا .

الا أن هذا النقد لا يقلل من شأن الاختبارات الجيدة الصنع المتقنة البناء ، وقد كانت له تضميناته الايجابية ، فأصبح ناشرو الاختبارات في وقتنا الحاضر يقدمون لمستخدميها بيانات تحليلية تفصيلية عن الأسئلة التي يتألف منها الاختبار مع تحليل احصائي ومنطقي لها . كما أن هذا النقد جعل مستخدمي الاختبارات أكثر حرصا وحذرا في استخدامها بالاضافة الى شعورهم بالحاجة الى عدم الاعتماد على نتائج الاختبارات وحدها ، وانما لابد من تدعيمها بمعلومات اضافية من المصادر الأخرى .

ويرتبط بهذا النقد نقد آخر يتصل بمدى شمول الاختبارات لنطاق واسع من السلوك الانساني . اننا نعلم أن الاختبارات لا تقيس في معظم الحالات الا عينات من السلوك الانساني (اذا تجاوزنا الاختبارات المرجعة الى المحك ذات الاستخدامات الخاصة في مجال التعليم والتدريب) ، ومع ذلك نقع في الخطأ الشائع الذي يتمثل في أننا نتوقع منها ما لم تقصد اليه . ولعلنا نكون أكثر علمية ودقة اذا استخدمنا الاختبارات في تحقيق الوظائف المحددة التي تسعى اليها ، انها حينئذ تصبح أدوات جديدة للعلم . ولعل هذا القول يعود بالنقاد الذين يوجهون الى اختبارات الاستعدادات ومنها اختبارات الذكاء في وقتنا الحاضر عاصفة من الهجوم الى جادة الصواب . ان اختبار الاستعداد الذي يستخدم في أغراض الانتقاء لمؤسسة تعليمية أو مهنية مثلا لا يحقق الغرض منه اذا افترض فيه كل من الناقد والمدافع جميعا أنه يقيس جميع المتطلبات اللازمة للنجاح في دراسة معينة أو مهنة معينة ، فهذا القول يتجاوز المعطيات والبيانات التي توفرها الاختبارات . وحتى لا نقع في أي تطرف جديد علينا دائما أن نتذكر أن الدرجات التي نحصل عليها من الاختبارات ليست الا نوعا من المعلومات - من بين أنواع عديدة أخرى متاحة للاخصائي النفسي ، ولا يجب عليه أن يصدر حكمه أو يتخذ قراره بناء على نتائج الاختبارات وحدها .

ان الاختبارات النفسية - كغيرها من وسائل أدوات البحث - قد تختلف في درجة الدقة ، وعلينا الا نتوقع منها ما ليس من طبيعتها ، وعلينا ان نجعل نصب أعيننا دائما انه اذا كانت المقاييس الفيزيائية - وما هي عليه من سمعة وتاريخ - تعاني من اخطاء القياس واططاء الملاحظة واططاء التنبؤ ، فما بالك بالاختبارات النفسية وهي مقاييس للسلوك الانساني بكل ما فيه من تعقد وتغير ورقى .

٢ - اثر الاختبارات في المفحوص :

يرجى النقد الى الاختبارات نقدا آخر يتلخص فيما تحسنته هذه الاختبارات من آثار غير مرغوب فيها - ان لم تكن ضارة - في المفحوصين من مختلف المستويات العمرية . ولعل أشهر هذه الآثار ذكرا هو القلق . وبالطبع لا ينكر أحد ان الموقف الاختباري يحدث في المفحوص قدرا من القلق ، الا انه من الحقائق السيكولوجية الهامة ان قدرا معقولا من القلق يدفع المفحوصين الى مزيد من التحسن في الأداء . وقد يكون الاثر الضار للقلق مرتبطا بالمفحوصين ذوي المستويات العليا فيه حيث يتداخل القلق الزائد مع الأداء الاختباري . وقد لا يكون في ذلك مسئولية مباشرة للاختبار ذاته ، فكل ما يفعله القلق مع هؤلاء المفحوصين ذرى القلق المرتفع ان يضيف اليه المزيد . ويوجد في فن تطبيق الاختبارات قواعد اساسية لو القزم بها الفاحصون يمكنهم خفض القلق الاختباري ، ويمكن ان يتم ذلك في جميع مراحل تطبيق الاختبار ، سواء عند التهيؤ له ، او التدريب على الأمثلة التوضيحية او استخدام الوقت الحد في الأداء الاختباري (اختبارات القوة مثلا) . ولحسن الحظ يتحول ناشرو الاختبارات في الوقت الحاضر الى تزويد المفحوصين باختبارات يمكنهم التدريب عليها وتوافر في مكتبات الجامعات الأمريكية والأوربية أمثلة كثيرة من هذه الاختبارات التدريبية ، وبالإضافة الى ذلك فان مؤلفي اختبارات الأداء الاقصى يتحولون ايضا في الوقت الحاضر الى التركيز على اختبارات القوة (التي لا تتحدد بوقت معين) بدلا من اختبارات السرعة . وفي جميع الأحوال لابد للفاحص من ملاحظة المفحوص بعمق أثناء الموقف الاختباري ، وإعادة النظر في درجات المفحوصين ذوي الانفعالية العالية وخاصة ان قلق الاختبار ليس له الا اثر مؤقت في الصحة

النفسية للمفحوص • ومرة أخرى فإن المسألة تعتمد على الاستخدام الجيد والتوظيف الرشيد للاختبار النفسي ونتائجه •

٣ - الأثر الاجتماعي والثقافي للاختبارات النفسية :

لعل هذه المسائل هي أخطر المسائل على الإطلاق ، وتأتي أهميتها من أنها شغلت الباحثين في مجال التقويم النفسي وخارجه سنوات طويلة تكاد تتساوى مع العمر الزمني لحركة القياس النفسي ذاتها • فكثيرا ما يقال ان الاختبارات النفسية غير عادلة في تعاملها مع الفئات الاجتماعية داخل المجتمع ومن السلالات البشرية بين المجتمعات المختلفة • وقد تزايدت هذه الانتقادات في السنوات الأخيرة في بعض المجتمعات التي مورست فيها الاختبارات النفسية لسنوات طويلة ، وعلى الأخص بريطانيا والولايات المتحدة الأمريكية ، ويرجع ذلك الى زيادة المطالبة بالحقوق المدنية وتهيئة تكافؤ الفرص للجميع •

وتتلخص الاتهامات الاجتماعية التي يوجهها النقاد الى الاختبارات النفسية بأنها متحيزة اجتماعيا وتمثل حواجز ضد الفرص التعليمية والمهنية لمعدي من أبناء الفئات المحرومة داخل المجتمع الواحد •

وحتى يمكن لنا مناقشة هذه الاتهامات بشيء من الموضوعية نذكر أن مسألة التحيز أو عدم العدالة في الاختبار تشير الى مسألتين مختلفتين ، أولاهما تتصل ببناء الاختبار ذاته حيث يتألف من أسئلة أو مفردات تلائم فئات اجتماعية دون سواها ، وثانيتهما تتصل بالاستخدام الظالم لنتائج الاختبارات وخاصة فيما يتصل باصدار الأحكام أو اتخاذ القرارات ، وهي مسألة تخرج بالباحث من نطاق العلم الى التحيز الصريح •

أما فيما يخص المسألة الأولى فأننا نجد ان معظم ما يكتب عن تحيز محتوى الاختبار يتركز على أداء بعض الجماعات أو الأفراد في الاختبار يكون متخلفا ، فيحصلون على درجات منخفضة بالمقارنة بغيرهم من الجماعات أو الأفراد • الا أن هذا ليس حجة ويتجاوز الدليل الموضوعي ، فالأداء المنخفض في ذاته ليس برهانا كافيا على تحيز الاختبار ، والا كان كل اختبار في الحساب يصبح متحيزا ضد الضعاف في الرياضيات ، وكل

اختبار فى المفردات اللغوية يصبح متحيزا ضد أصحاب الحصول اللغوى المتواضع ، ان هذا القول لا يصبح صحيحا الا اذا تدخلت الأفراد فى الاختبار عوامل أخرى لم يتحكم فيها الباحث ، كأن تصاغ المشكلات الحسابية فى قالب لغوى يحتاج الى قدرة لغوية عالية ، أما اذا تحرر الاختبار الحسابى من هذه العوامل (أى كان صادقا بالمعنى الدقيق للكلمة) فلماذا لا يكون ضعف الأداء فيه دليلا على نقص فى القدرات والمهارات الحسابية تحتاج الى نوع من التربية التعويضية أو التعلم العلاجى .

الا ان ما نحب أن ننبه اليه هو أن بعض الاختبارات المقننة تركز على المحتوى والقيم الثقافية المألوفة والشائعة لدى الانسان الغربى الأبيض من الطبقة المتوسطة (فقد كانت الاختبارات منذ نشأتها الاولى غربية المنشأ) قالب لغوى يحتاج الى قدرة لغوية عالية ، أما اذا تحرر الاختبار الحسابى والسلالات البشرية الأخرى . ولذلك لابد من أن يتنبه الباحثون فى الثقافات الأخرى الى عدم اللجوء الى « النقل الأعمى » أو « الترجمة الحرفية » لهذه الاختبارات ، والا أوقعوا أنفسهم وثقافتهم فى مأزق اجتماعى خطير . ولعلنا بهذا نشير الى بعض الممارسات الخاطئة فى حركة الاختبارات النفسية فى الثقافتين المصرية والعربية حيث ينطبق على معظمها خصائص النقل الأعمى والترجمة الحرفية . بل اننا بهذا ندعو الى ابداع ثقافى جديد فى مجال الاختبارات النفسية ، فليس كل اختبار ظهر فى عاصمة أوروبية أو فى جامعة أمريكية قابل للنقل ، فبعضه وراءه من الافتراضات النظرية فى علم النفس والأسس الثقافية للمجتمع ما يجعل نقله - حتى ولو كان مع بعض التعديل - نوعا من « المجزرة » العقلية كما قال - بحق - عالم النفس التركى جوندوز فاسف فى بحثه عن الاختبارات النفسية فى العالم الثالث والذى ألقاه فى المؤتمر الدولى الثالث والعشرين لعلم النفس الذى انعقد بالمكسيك عام ١٩٨٤ .

الفصل الثانى

مدخل الى الاحصاء النفسى

- أشرنا فى الفصل السابق الى أن القياس النفسى يعتمد فى جوهره على البيانات الكمية التى تتوافر عن السلوك الانسانى ، وهى ، كالبيانات الكيفية يجب أن تتاح للاحصائى النفسى كى يصدر أحكامه على هذا السلوك ويتخذ القرارات العملية حوله .

وبالطبع فإن البيانات الكمية التى تستند اليها عملية التقويم النفسى قد تكون أدق من البيانات الكيفية ، ويتوقف ذلك على طبيعة الظاهرة النفسية موضع البحث والدراسة ، ويجب على الأخصائى الا يتعسف فيفرض على هذه الظاهرة طبيعة غير طبيعتها ، ومن ذلك أننا حين نقارن بين شخصين فى « طول القامة » فإن من الأفضل أن نصف الشخص الأول مثلا بأن طوله ١٦٠ سم ونصف الشخص الآخر بأن طوله ١٨٠ بدلا من تقول ان الشخص الثانى أطول من الأول ، فالعبارة الأولى استخدمت لغة الكم ، وهى أقرب الى المقام والى طبيعة الظاهرة المقيسة ، ولكننا عند المقارنة من شخصين فى القدرة على القيادة ، يكون وصفنا اللفظى المفصل المعتمد على بصيرة الباحث أجدى وأفضل .

وتحتاج البيانات التى يحصل عليها الباحث الى تحليل ، ويستخدم علم الاحصاء فى تحليل البيانات الكمية خاصة أما البيانات الكيفية فلها طرقها الخاصة .

ونقدم فى هذا الفصل عرضا موجزا لطرق تحليل البيانات الكمية مصنفة تبعا لمستوى المقياس المستخدم . وقد سبق أن أشرنا الى أن مستويات القياس هذه أربعة وهى المقاييس الاسمية ، ومقاييس الرتبة ومقاييس المسافة ومقاييس النسبة وسوف نقتصر على الطرق المستخدمة فى الاحصاء الوصفى خاصة كما يتمثل فى الاتجاهات الرئيسية الثلاثة ، التزعة المركزية والتشتت والعلاقة، مع ملاحظة أن الطرق الاحصائية التى تصلح للمستويات الدنيا تصلح أيضا للمستويات العليا والعكس غير صحيح .

تحليل بيانات المقاييس الاسمية

أشرنا الى أن المقاييس الاسمية تستخدم الأعداد لتشير الى أفراد أو هئات ، وفي حالة استخدام الفئات ، فاننا نقوم « بعد » عدد الحالات التي تقع في كل فئة ، ويطلق على هذا العدد « التكرار » وتوجد طرق احصائية عديدة للتعامل مع التكرارات المصنفة الى فئات ، اما الاعداد التي تشير الى الفئات ذاتها فلا تخضع لأي تحليل احصائي لأنها في الواقع محض رموز . وقد يشير المقياس الاسمي الى ما يسمى القيم المنفصلة وفي هذه الحالة ينطبق على أسئلة الاختبار اذا كانت الاجابة عليها من نوع « الكل أو لا شيء » ، مثل « اجابة صحيحة أو خاطئة » أو « نعم أولا » وعملت منفصلة ، أي كل سؤال على حدة .

ويحتاج الباحث كما اشرنا في تحليله لبيانات المقاييس الى استخدام الطرق الاحصائية في تحديد الاتجاهات الرئيسية الثلاثة : النزعة المركزية ، والتشتت ، والعلاقة . وقد لا تتضح هذه الاتجاهات جميعا في المقاييس الاسمية ، بينما نجدها كلها في مقاييس الرتبة ، ومقاييس المسافة والنسبة كما سوف تلاحظ .

النسب والنسب المئوية :

عادة ما يستخدم في تحليل التكرارات النسب والنسب المئوية ، والمثال التالي يوضح ذلك :

لاحظ احد المدرسين اختلاف عدد التلاميذ والتلميذات في فصلين من فصوله ، كما يتبين من الجدول التالي :

تكرار التلاميذ	تكرار التلميذات	المجموع
٣٠	٢٠	٦٠
٣٠	٢٠	٥٠
الفصل (١)		
الفصل (ب)		

انه لا يستطيع أن يستنتج من الجدول السابق ان الوضع النسبى للذكور فى الفصلين متساو مادام عدد التلاميذ الذكور متساو ، ويفيد فى المقارنة ان يقوم بتحويل التكرارات المتضمنة فى الجدول السابق الى نسب او نسب مئوية كما يلى :

نسب التلاميذ	نسب التلميذات	النسبة المئوية للتلاميذ	النسبة المئوية للتلميذات
٥٠	٥٠	%٥٠	%٥٠
٦٠	٤٠	%٦٠	%٤٠

وقد حصلنا على نسبة الذكور مثلا بقسمة تكرارهم على المجموع الكلى لتلاميذ الفصل (اى $\frac{50}{100}$) اما النسبة المئوية فقد حصلنا عليها بضرب النسبة السابقة $\times 100$ ، ولعلك لاحظت بعد هذا التحويل ان الذكور يمثلون نسبة ٥٠ (نسبة مئوية) اكبر من الاناث فى الفصل (ب) ، بينما تقسأوى النسبتان فى الفصل (ا) وتدل بالطبع النسبة ٥٠ او النسبة المئوية ٥٠% على نقطة التوسط او النزعة المركزية . وقد يستخدم فى قياس النزعة المنوال mode ويدل على اكثر القيم شيوعا فى التوزيع التكرارى ، وعلى الرغم من أنه مقياس مباشر للنزعة المركزية الا أنه اقل دقة من المتوسط والوسيط اللذين سنشير اليهما فما بعد .

المدى المطلق :

المدى المطلق هو أحد مقاييس التشتت واكثر بساطة وسهولة ، وهو فى جوهره مقياس للسعة ، ويحسب مباشرة بتحديد الفرق بين اكبر عدد واقل عدد فى البيانات المتوافرة ، وهو بهذا المعنى يصلح للاستخدام مع المقاييس الاسمية ، فاذا كانت ارقام جلوس التلاميذ تمتد من ١ الى ٥٠٠ فان المدى المطلق فى هذه الحالة يصبح ٤٩٩ ، ويمكن استخدامه بنفس الطريقة مع درجات الاختبارات .

وقد يستخدم فى قياس تشتت بيانات المقاييس الاسمية والنسب

المئوية أيضا . فكلما ازدادت هذه النسب تطرقا بالزيادة أو النقص دل ذلك على درجة البعد عن نقطة التوسط (٥٠ %) .

معامل الارتباط بين بيانات المقاييس الاسمية :

يدل معامل الارتباط على مدى الاتفاق بين متغيرين ، أى ما اذا كانا يختلفان معها ، وفى أى اتجاه ، وإلى أى حد ؟ أو معنى آخر هو النزعة الى اقتران التغير فى ظاهرة بالتغير فى ظاهرة أخرى .

يتمتع معامل الارتباط بين (+ ١) والذي يعنى وجود ارتباط موجب كامل بين المتغيرين ، (فالصفر) الذى يعنى عدم وجود ارتباط بينهما واستقلالهما عن بعضهما تماما ، (- ١) الذى يعنى وجود ارتباط سالب كامل بين المتغيرين ، والواقع ان الحصول على معاملات الارتباط الكاملة ، سواء بالموجب أو السالب ، نادر الحدوث ان لم يكن مستحيلا . ولذلك فان اغلب معاملات الارتباط بين المتغيرات فى المجالات النفسية والتربوية والاجتماعية عبارة عن كسر عشرى (اكبر من الصفر وأقل من الواحد للصحیح) يدل لى وجود علاقة جزئية ، وقد يكون هذا الكسر العشرى موجبا أو سالبا .

وتوجد طرق احصائية كثيرة لقياس العلاقة بين التكرارات المصنفة الى فئات المقاييس الاسمية ، وهى تنتمى الى ما يسمى الاحصاء البارامترى ، ولا يتسع المقام لعرض هذه الطرق ويمكن الرجوع فى ذلك الى مرجع متخصص (٧) ، ونكتفى بعرض طريقة واحدة هى ما يسمى عامل ارتباط فاي .

مثال : حصل احد المعلمين على البيانات التالية من عدة فصول يقوم بالتدريس لها تتألف من الجنسين (الذكور والاناث) صنفوا جميعا فى ضوء نجاحهم أو فشلهم الدراسى كما يلى :

التلاميذ	التلميذات	المجموع	
(٣٥)	(٣٧)	٧٢	ناجحون
(١٤)	(٣٤)	٤٨	فاشلون
٤٩	٧١	١٢٠	المجموع

واراد هذا المعلم أن يحلل البيانات الاسمية فى الجدول السابق ليستخلص منه ما اذا كانت توجد علاقة بين الجنس والنجاح المدرسى . انه عليه فى هذه الحالة أن يحسب معامل ارتباط فاي ، وذلك بتحويل التكرار الثنائى للاعداد بين القوسين فى الجدول السابق الى اللغة الرمزية كما يلى :

التلاميذ	التلميذات
١	ب
ج	د

ثم تطبق المعادلة الاتية :

$$R = \frac{ad - bc}{(a+b)(c+d)(a+c)(b+d)}$$

وبالتعويض عن رمز المعادلة فان :

$$R = \frac{(14 \times 37) - (34 \times 35)}{(71 \times 59 \times 48 \times 72)} + 19$$

وهو معامل ارتباط موجب ولكنه منخفض ، وقد لا تكون له دلالة احصائية (*) اى لا يدل على وجود علاقة بين الجنس والنجاح المدرسى .

تحليل بيانات مقياس الرتبة

يمكن الحصول على اتجاه النزعة المركزية للبيانات الرتبية باستخدام الوسيط ، وعلى اتجاه التشتت باستخدام نصف المدى الربيعى ، وعلى معامل الارتباط كمقياس للعلاقة باستخدام معامل ارتباط الرتب .

الوسيط :

يمكن تعريف الوسيط بأنه الرتبة التى تقسم التكرارات الى نصفين متساوين ، وفيما يلى مثال يوضح طريقة حسابه :

فئات الدرجات	الحدود الحقيقية للفئات	التكرار	التكرار المتجمع الصاعد
٥ - ٩	٤ر٥ - ٩ر٥	٤	٤
١٠ - ١٤	٩ر٥ - ١٤ر٥	١	٥
١٥ - ١٩	١٤ر٥ - ١٩ر٥	١	٦
٢٠ - ٢٤	١٩ر٥ - ٢٤ر٥	١٠	١٦ فئات الوسيط
٢٥ - ٢٩	٢٤ر٥ - ٢٩ر٥	٣	١٩
٣٠ - ٣٤	٢٩ر٥ - ٣٤ر٥	٥	٢٤
٣٥ - ٣٩	٣٤ر٥ - ٣٩ر٥	٣	٢٧
٤٠ - ٤٤	٣٩ر٥ - ٤٤ر٥	٠	٢٧
٤٥ - ٤٩	٤٤ر٥ - ٤٩ر٥	١	٢٨
		الجموع	٢٨

(*) مفهوم الدلالة الاحصائية ينتمى الى الاحصاء الاستدلالى ، ولا يتسع المقام فى هذا الكتاب لشرحه ويمكن للقارئ الرجوع الى بعض المصادر المتصلة فى الاحصاء النفسى والتربوى والاجتماعى .

ولحساب الوسيط من بيانات الجدول السابق يتبع الباحث الخطوات الآتية :

١ - تحديد رتبة الوسيط أى النقطة التى ينقسم عندها التكرار إلى نصفين بقسمة المجموع الكلى للتكرارات على ٢ أى أن رتبة الوسيط

$$28 \\ 14 = \frac{\quad}{2} =$$

٢ - دراسة التكرارات المتجمعة الصاعدة (العمود الأخير فى الجدول) لتحديد الفئة التى يقع فيها الوسيط ، وهى فى هذه الحالة الفئة (٢٠ - ٢٤) .

٣ - تطبيق المعادلة التالية :

$$\text{الوسيط} = \text{الحد الأدنى الحقيقى لفئة الوسيط} + \left[\frac{\text{رتبة الوسيط} - \text{التكرار السابق لفئة الوسيط}}{\text{تكرار فئة الوسيط}} \right] \times \text{مدى الفئة}$$

على النحو الآتى :

$$\text{الوسيط} = ١٩٥ + ٥ \times \frac{١٤ - ٦}{١٠} = ٢٢٥$$

نصف المدى الربيعى :

لعلك لاحظت أن الوسيط هو نقطة واحدة تقسم التكرار الكلى أى مجموع الأفراد أو الحالات إلى نصفين ، ويمكن لهذه الفكرة أن تمتد منطقياً إلى تقسيم التكرار إلى أى عدد من الأقسام ، فمثلاً يمكن أن ينقسم إلى أربعة أقسام باستخدام ثلاث نقاط التى تسمى فى هذه الحالة الربيعيات ، أو إلى عشرة أقسام باستخدام تسع نقاط تسمى العشريات ، أو مائة قسم باستخدام ٩٩ نقطة تسمى المئنيات ، وتلتقى هنا بفكرة الربيعيات التى يعتمد عليها حساب نصف المدى الربيعى كمقياس إحصائى لتشتت البيانات الرتبىة .

ولا يختلف حساب الربيع عن حساب الوسيط ، إلا في نقطة البداية وعلى هذا يمكن حساب الربيع الأدنى أو الأول من بيانات الجدول السابق كما يلي :

١ - تحديد رتبة الربيع الأول بقسمة المجموع الكلى للتكرار على (٤) وضربها في (١) أى أن :

$$\text{رتبة الربيع الأول أو الأدنى} = \frac{1 \times 28}{4} = 7$$

وهم ربع الأفراد الذين يفترض فيهم أن تقع درجاتهم أدنى من نقطة الربيع الأول أو الأدنى ، أما الأرباع الثلاثة الأخرى فتقع درجاتهم أعلى منها :

٢ - دراسة التكرارات المتجمعة الصاعدة لتحديد الفئة التى تقع فيها نقطة الربيع الأول وهى هذه الحالة الفئة (٢٠ - ٢٤) .

٣ - تطبيق المعادلة التالية التى لا تختلف فى صيغتها عن معادلة حساب الوسيط .

$$\text{الربيع الأول} = \text{الحد الأدنى الحقيقى لفئة الربيع الأول} + \left[\frac{\text{رتبة الربيع الأول} - \text{التكرار المتجمع السابق لفئته}}{\text{تكرار فئته}} \right] \times \text{مدى الفئة}$$

أى أن :

$$\text{الربيع الأول} = ١٩ + ٥ \times \frac{٧ - ٦}{١٠} = ٢٠$$

أما الربيع الثانى فتحدد من رتبته على النحو التالى بافتراض ن = المجموع الكلى للتكرار .

$$\text{رتبة الربيع الثانية أو الأوسط} = \frac{ن}{2} = \frac{ن \times ٢}{4} = \text{الوسيط}$$

وهم نصف الأفراد الذين يفترض فيهم أن تقع درجاتهم أدنى من نقطة الربيع الثانى أو الوسيط ، وبالمطبع فإن النصف الثانى تقع درجاتهم أعلى منها .
(التقويم النفسى)

ومعنى ذلك : ان الربيع الثانى هو الرميض وسبق لنا حسابه .

اما الربيع الثالث او الاعلى فيحدد من بيانات الجدول السابق كمايلى :

$$١ - رتبة الربيع الثالث او الاعلى = \frac{ن \times ٣}{٤} = \frac{٢٨ \times ٣}{٤} = ٢١$$

وهم ثلاثة ارباع الأفراد الذين يفترض فيهم أن تقع درجاتهم ادنى من نقطة الربيع الثالث او الاعلى ، اما الربيع الباقي فتقع درجاتهم اعلى منها

٢ - الفئة التى يقع فيها الربيع الثالث هى ٣٠ - ٣٤

$$٢ - الربيع الثالث = ٢٩٥ + \left(\frac{٢١ - ١٩}{٥} \right) \times ٥ = ٣١٥$$

ولحساب نصف المدى الربيعى كمقياس احصائى لتشتت البيانات الرتبية تعتمد على نصف الفرق بين الربيعيين الاول والثالث وهو ما يسعى نصف المدى الربيعى .

نصف المدى الربيعى =

الربيع الثالث - الربيع الاول

٢

$$= \frac{٣١٥ - ٢٠}{٢} = ١٤٧٥$$

معامل ارتباط الرتب :

يستخدم فى القياس الاحصائى للعلاقة بين متغيرين من طبيعة رتبية

معامل ارتباط الرتب الذى ابتكره عالم النفس الانجليزى تشارلز سبيرمان ،
ويوضح الجدول التالى طريقة حسابه .

مثال :

حصل أحد الباحثين على ترتيب ١٥ تلميذا فى مادتي التاريخ
والجغرافيا . حيث تدل الرتبة (١) على التلميذ الأول والرتبة (١٥) على
التلميذ الأخير فى كل مادة على حدة ، مع ملاحظة أن التلاميذ الذين تساوت
رتبتهم حصلوا على متوسط الرتب المتتالية ، أى ان التلاميذ د ، س حصلوا فى
التاريخ على رتبة ١٢ر٥ وهى متوسط الرتبتين ١٢ ، ١٣ ، والتلاميذ ا ، ز ،
ن ، حصلوا جميعا على الرتبة (٨) وهى متوسط الرتب ٧ ، ٨ ، ٩ .

التلاميذ	ترتيب التاريخ	ترتيب الجغرافيا	فروق الرتب (ف)	مربع فروق الرتب (ف²)	
ا	١١	٨	+	٣	٩ر٠٠
ب	٤	٦	-	٢	٤ر٠٠
ج	٩	٥	+	٤	١٦ر٠٠
د	١٠	١٤	-	٤	١٦ر٠٠
هـ	١٤ر٥	١٥	-	٠ر٥	٠ر٢٥
و	١٤ر٥	١٢	+	٢ر٥	٦ر٢٥
ز	١٢ر٥	٨	+	٤ر٥	٢٠ر٢٥
ح	١	٣	-	٢	٤ر٠٠
ط	٣	١	+	٢	٤ر٠٠
ى	٧	٤	+	٣	٩ر٠٠
ك	٦	١٠	-	٤	١٦ر٠٠
ل	٢	٢		صفر	٠٠ر٠٠
م	٥	١٣	-	٨	٦٤ر٠٠
ن	٨	٨		صفر	٠٠ر٠٠
س	١٢ر٥	١١	+	١ر٥	٢ر٢٥

مجموع ف = صفر = صفر مجموع ف٢ = ١٧١ر٠٠

ويجيب معامل الارتباط الرقب بالمعادلة الآتية :

$$r = 1 - \left[\frac{6 \times \text{مجموع } F^2}{n(n^2 - 1)} \right]$$

$$r = 1 - \frac{171 \times 6}{234 \times 15} = 0.696$$

وهو معامل ارتباط موجب مرتفع نسبيا ، ويدل على وجود علاقة بين التحصيل في كل من التاريخ والجغرافيا أى أن التلميذ الذى جاء ترتيبه مرتفعا في التاريخ يحتمل أن يكون كذلك في الجغرافيا والعكس صحيح .

تحليل بيانات مقاييس المسافة والنسبة

يحصل الباحث على اتجاه النزعة المركزية للبيانات التى توفرها مقاييس المسافة والنسبة باستخدام المتوسط الحسابى ، وعلى اتجاه التشتت بحساب الانحراف المعياري ، وعلى معامل الارتباط باستخدام طريقة بيرسون التى تسمى طريقة الارتباط التتابعى ، (أو طريقة حاصل ضرب العزوم) .

المتوسط الحسابى :

المتوسط الحسابى كمقياس للنزعة المركزية لبيانات مقاييس المسافة والنسبة ، يعتمد على تقسيم الدرجات أو الكميات - وليس التكرارات كما هو الحال في الوسيط - الى نصفين متساويين ، ويوضح المثال التالى ذلك :

مثال :

طبق أحد المعلمين اختبارا في فهم القراءة على تلاميذ أحد فصول الصف الرابع الابتدائى فحصل على البيانات الآتية :

التلاميذ	أ	ب	ج	د	هـ	و	ز	ح	ط
الدرجات	١	٢	٤	٥	٧	٥	٣	١	٧

وللحصول على المتوسط الحسابي كمقياس احصائي للنزعة المركزية يقوم المعلم بجمع درجات التلاميذ ، وهو في هذه الحالة ٢٦ ، وجمع عدد التلاميذ وهو في هذه الحالة ٩ ، ويحسب المتوسط الحسابي بالمعادلة الآتية :

$$\text{المتوسط الحسابي} = \frac{\text{مجموع الدرجات}}{\text{عدد الأفراد}} = \frac{\text{مجموع } 26}{9} = \frac{26}{9} = 2.88 \approx 3$$

الانحراف المعياري :

يعتمد حساب الانحراف المعياري كمقياس احصائي لتشتت بيانات مقاييس المسافة والنسبة على حساب المتوسط الحسابي أولاً ، ثم حساب الفروق أو الانحرافات عن هذا المتوسط ، ونظراً لأن مجموع هذه الفروق يساوي صفراً ، يستحيل علينا أن نحسب متوسط هذه الفروق ، ولذلك يلجأ الاحصائيون إلى تربيع كل فرق أو (انحراف) للتخلص من العلامات الجبرية السالبة ، ثم نحسب بعد ذلك متوسط مربعات الانحرافات بقسمتها على عدد الأفراد ، ثم نستخرج الجذر التربيعي لمتوسط مربعات الانحرافات ، ويكون هذا الجذر التربيعي هو الانحراف المعياري .

ويمكن حساب الانحراف المعياري للمثال السابق ، وذلك بطرح المتوسط الحسابي ($\bar{x} = 3$) ، من كل درجة من الدرجات السابقة فنحصل على الانحرافات عن المتوسط (ح) ، ثم نربعها (ح^٢) كما يلي :

التلاميذ	أ	ب	ج	د	هـ	و	ز	ح	ط
ح	٢-	١-	صفر	١+	٢+	١+	١-	٢-	٢+
ح ^٢	٩	١	صفر	١	٩	١	١	٩	٩

مجموع مربعات الانحرافات = ٤٠
ثم نحسب الانحراف المعياري بالمعادلة التالية :

$$\text{الانحراف المعياري} = \sqrt{\frac{\text{مجموع مربعات الانحرافات عن المتوسط}}{\text{عدد الأفراد}}}$$

$$\sigma = \sqrt{\frac{\sum x^2}{n}} = \sqrt{\frac{40}{11}} = 1.91$$

معامل الارتباط التتابعي لبيرسون :

يستخدم في حساب العلاقة بين مقاييس المسافة والنسبة الطريقة العامة في حساب معامل الارتباط والتي ابتكرها عالم الاحصاء الانجليزي كارل بيرسون ، ويوضح المثال التالي طريقة الحصول على هذا المعامل .

مثال : طبق أحد المعلمين اختبارين يقيس احدهما تحصيل تلاميذ الصف الرابع من التعليم الاساسي في مادة الحساب (س) ويقيس الاخر القدرة العددية (هـ) لدى هؤلاء التلاميذ ، ويوضح الجدول الآتي البيانات التي حصل عليها وخطوات حساب معامل الارتباط .

التلاميذ	درجات اختبار التحصيل الحسابي (س)	درجات اختبار القدرة العددية (هـ)	س	هـ
أ	٢	٥	٤	٢٥
ب	٢	٧	٩	٤٩
ج	٥	٦	٢٥	٢٦
د	٧	١٠	٤٩	١٠٠
هـ	٨	١٢	٦٤	١٤٤
ن = ٥	مجموع س = ٢٥	مجموع هـ = ٤٠	مجموع س ^٢ = ١٥١	مجموع هـ ^٢ = ٥٢٤
			مجموع س هـ = ٢٢٧	

ويمكن حساب معامل الارتباط مباشرة من بيانات الجدول السابق
بتطبيق المعادلة الآتية :

$$r = \frac{(n \text{ مج س ص}) - (\text{مج ص} \times \text{مج س})}{\sqrt{[(n \text{ مج س}^2) - (\text{مج س})^2] \times [(n \text{ مج ص}^2) - (\text{مج ص})^2]}}$$

وبالتعويض عن رموز المعادلة السابقة نجد أن :

$$r = \frac{(227 \times 5) - (40 \times 25)}{\sqrt{[(151 \times 5) - 625] \times [354 \times 5 - 1600]}}$$

وهو معامل ارتباط مرتفع جداً ، ويدل على أن التلميذ الذى يحصل
على درجات عالية فى اختبار التحصيل الحسابى يحتمل احتمالاً كبيراً أن
يكون على درجة عالية فى اختبار القدرة العددية ، والعكس صحيح .

مراجعة الباب الأول

١. - أحمد زكي صالح : الأسس النفسية للتعليم الثانوي : القاهرة : دار النهضة العربية ، ١٩٧٢ .
٢. - أحمد زكي صالح : علم النفس التربوي : القاهرة : النهضة المصرية ، ١٩٧٢ .
٣. - جابر عبد الحميد : التقويم التربوي والقياس النفسي . دار النهضة العربية ، ١٩٨٣ .
٤. - سعيد يامشعوس ، السيد محمد خيرى ، يحيى مهنى : التقوية التربوي الرياض : منشورات دار الفصيل الثقافية ، ١٩٨٠ .
٥. - صلاح الدين محمود علام : تحليل البيانات فى البحوث النفسية والتربوية . دار الفكر العربى ، ١٩٨٥ .
٦. - رايتسون ، ج . جاستمان ، ج . ر . روبنز ، (ترجمة : محمد محمد عاشر وآخرين) : التقويم فى التربية الحديثة . القاهرة : الأنجلو المصرية ، ١٩٦٥ .
٧. - رمزية الغريب : القياس اللابرمترى . مكتبة الأنجلو المصرية ، ١٩٨٥ .
٨. - رمزية الغريب : التقويم والقياس النفسى والتربوي . القاهرة : الأنجلو المصرية ، ١٩٧٠ .
٩. - عبد الله السيد عبد الجواد : المؤشرات التربوية واستخدام الرياضيات فى العلوم الانسانية ، أسبوط ، ١٩٨٣ .
١٠. - فؤاد أبو حطب . القدرات العقلية . القاهرة : الأنجلو المصرية ط ٤ ، ١٩٨٣ .
١١. - فؤاد أبو حطب ، آمال صادق ، علم النفس التربوي ، الأنجلو المصرية ، (ط ٢) ، ١٩٨٤ .
١٢. - فؤاد أبو حطب ، بديوى علام : الاحصاء النفسى والتربوي ، مكتبة الأنجلو المصرية (تحت الطبع) .
١٣. - فؤاد البهى السيد : علم النفس الاحصائى وقياس العقل البشرى . دار الفكر العربى ، ١٩٧٨ .
١٣. - محمد خليفة الترنسى : التقويم والتقييم ، مجلة العربى ، يونيه ١٩٨٦ .
١٣. ب - محمد خليفة الترنسى اضرأء على لغتنا السبعة ، كتاب العربى ، ٨٦ .
١٣. ج - محمد خليفة الترنسى : الاشتقاق من المشتق ، مجلة العربى ، يولية ٨٦ .
١٤. - محمد عبد السلام احمد : القياس النفسى والتربوي . القاهرة النهضة المصرية ، ١٩٦٠ .
١٥. - نعيم عطية : التقييم التربوي الهادف ، بيروت : منشورات دار الكتاب اللبنانى ، ١٩٧٠ .

١٦ - يوسف محمود الشيخ • جابر عبد الحميد جابر : سيكولوجية الفروق الفردية • القاهرة : النهضة العربية ١٩٦٤ •

17. Allport, G.W. Pattern and growth in personality, New York: Holt Rinehart, 1961, (1937).
18. Anastasi, A. Psychological testing New York, Macmillan Publishing, (5th ed.) 1982.
19. Anastasi, A. Differential psychology (3rd ed.), New York : Macmillan, 1953.
20. Bloom, B. Taxonomy of educational objectives. London : Longmans, 1956.
21. Bordon, E.S. Psychological counseling. New York : Appleton, 1955.
22. Cook, W.W. The Function of measurement in the facilitation of Learning. In Linquist, E.F. (ed.) Educational Measurement Amer. Council on Educ., Washington, D.C., 1951, 3-64.
23. Coombs, C.H. Theory and method of social measurement. In : Festinger, L. and Katz (eds.) Research methods of behavioral Sciences. New Delhi Ed., 1970.
24. Cronbach, J. L. Essentials of psychological testing, New York : Harper (4th ed.), 1984.
25. Gronlund, Measurement and evaluation in teaching. New York : Macmillan, (4th ed.), 1981.
26. Krathwohi, D.R., et al. Taxonomy of educational objectives. II Affective domain, 1964.
27. Nunnally, J.C. (Jr.) Introduction to psychological measurement. New York : McGraw-Hill, 1970.
28. Rogers, C.R. Counseling and psychotherapy. Cambridge, Mass. : Riverside Press, 1942.
29. Rogers, C.R. Client-centered therapy. Cambridge, Mass.: Riverside Press, 1951.
30. Stevens, S.S. Mathematics, measurement and psychophysics. In : Stevens, S.S. (ed.) Handbook of Experimental Psychology. New York : Wiley, 1951.
31. Thorndike, R. and Hageu, E. Measurement and evaluation in psychology and education. New York : Wily, 1955.
32. Tylers, R.W. The Function of measurement in improving instruction. In: Lindquist, E.F. (ed.) Educational Measurement, 1951, 47-67.
33. Tyler, L.C. The work of the counselor (2nd ed.). New York York : Appleton-century, 1961.

الباب الثاني

شروط الاختبارات النفسية

الفصل الثالث

فن اعطاء الاختبارات

ماذا يقصد بتقنين الاختبار ؟

أشرنا في الفصل الأول الى تعريف للمقياس النفسى يشير الى ضرورة التعبير الصريح عن قواعده، وهذا ما يتضمنه مفهوم «التقنين» Standardization. فحين نقول ان هذا المقياس «مقنن» فان ذلك يعنى فى جوهره أنه لو استخدمه أفراد مختلفون يحصلون على نتائج متماثلة . ويتطلب هذا بالطبع توحيد اجراءات تطبيق الاختبار وتصحيحه ، فاذا كانت الدرجات التى يحصل عليها الأفراد المختلفون قابلة للمقارنة فان شروط الاختبار وظروف اعطائه يجب أن تكون موحدة للجميع . وهذا الشرط ما هو الا تطبيق خاص للحاجة الى الشروط والظروف المضبوطة فى جميع الملاحظات العلمية . ومثل هذا التقنين يجب أن يشمل المواد المستخدمة وحدود الزمن والتعليمات الشفوية والتحريرية التى تعطى للمفحوصين والأمثلة التوضيحية ، وطرق تناول أسئلة المفحوصين واستفساراتهم ، وكل ظرف أو حالة لا ترتبط بما يتناوله الاختبار ويمكن أن تؤثر فى أداء المفحوصين فى الاختبار كلما كان ذلك ممكنا .

ويوفر تقنين قواعد القياس للباحثين فى علم النفس ظروفا ملائمة لسهولة «الاتصال» ، لأن من أهم خصائص العلم وجود نوع من الاتصال الجيد بين العلماء بحيث يستطيع الباحث أن يقارن بين نتائجه ونتائج زملائه الذين يتناولون نفس المشكلات بالدراسة . فإذ افترضنا أن أحد الباحثين أجرى تجربة على آثار التوتر فى استجابات القلق وسجل نتائجه فى عبارة مثل أن «التوتر يؤدي الى ظهور القلق عند الأطفال» فان مثل هذه النتيجة لا تفيد فى «الاتصال» الجيد لأن الباحثين الآخرين قد يختلفون حول ما يقصده «بظهور القلق» بل قد يصعب عليهم أن لهم يستحل به تكرار نفس التجربة لتحقيق من صحة الفرض ، ويمكن تحقيق نوع من الاتصال الأكفأ لو احتطأ الباحث أن يلاحظ القلق باختبار «مقنن» أى محددة قواعده اعطائه وشروط تطبيقه .

وتتأثر عملية اعطاء الاختبار بعوامل متعددة ، منها ما يتصل بمن يعطى الاختبار وما يفعله ، ومنها ما يتصل بمن يقوم بإداء معين على هذا الاختبار وما يحيط به من ظروف ومؤثرات ، ومنها ما يتصل بالاختبار نفسه وما يتميز به من خصائص وما يقصد الى أن يقيس . وإذا كانت هناك عوامل وشروط ينبغي أن تتوفر في الاختبار نفسه حتى يكون أداة قياس يعتمد عليها ويطمأن إليها ، وهي ما سنتناولها في الفصول التالية ، فإن هناك شروطا وظروفا تؤلف ما نسميه فن اعطاء أو تطبيق الاختبارات على الفاحص أن يراعيها ويحققها ، وهي التي من شأنها أن توفر ، أول ما توفر ، الجو النفسي - الاجتماعي المناسب للمفحوص بحيث يأتي أدائه في الاختبار ، كما أراد واضعه ، وكما يريد الفاحص ، دقيقا وصادقا ، ليكون أساسا متينا لقرار تزيد احتمالات دقته وجدواه .

ولعل مسألة تقنين الاختبارات وغيرها من أدوات التقويم النفسي هي التي جعلت علماء النفس على درجة كبيرة من الحساسية لمسألة أخلاقيات المهنة . وقد ظهر هذا الاهتمام لأول مرة في أوائل الخمسينات حين أصدرت الجمعية الأمريكية لعلم النفس مجموعة من المعايير الأخلاقية التي يجب على الباحثين السيكولوجيين مراعاتها (١٢) ثم سرعان ما اتبعت ذلك بإصدار كتيب التوصيات الفنية الخاصة بالاختبارات النفسية والأساليب التشخيصية (١٣) ، وقد تطورت هذه الوثائق وتعدلت طوال السنوات الثلاثين الماضية وصدرت من كل منها طبعا جديدة (١٤ ، ١٥ ، ١٦ ، ١٧) . ونحن في حاجة الى ترشيد هذه الخبرة والممارسة السيكولوجيتين في مصر ، وقد عرض هذه الضرورة محمد خليفة بركات (١٠) في المؤتمر الثاني لعلم النفس في مصر عام ١٩٨٦ ، ولعل الجمعية المصرية للدراسات النفسية تقوم بهذه المسؤولية .

واجبات الفاحص ومسئوليته في اعطاء الاختبار :

١ - التدريب على اعطاء الاختبار : من الاختبارات ما يسهل اعطاؤه فلا يحتاج الى تدريب خاص للفاحص ، ومنها ما يحتاج الى تدريب ومران قد يمتد بضعة شهور . والاختبارات الفردية بصفة عامة تحتاج الى تدريب أكثر من الاختبارات الجماعية ، وإن كانت هناك بعض الاختبارات الجماعية التي تتطلب عناية كبيرة في اعطائها . وكذلك فإن الاختبارات الاسقاطية تحتاج

الى تدريب خاص قد يصل فى بعض الجامعات الى الحصول على دبلومة متخصصة فيها ، وهو ما لا تحتاجه الاختبارات او الاستبيانات ، وتحتاج المقابلة او الملاحظة تدريباً متخصصاً لا تحتاجه بعض وسائل التقرير الذاتى .
واذا كانت كل مسئولية الفاحص تنحصر فى قراءة مجموعة من التوجيهات والتعليمات المطبوعة ، فان أى شخص ذا ضمير يقظ حساس ، وليس فى شخصيته ما يحدث تهديداً للآخرين ، يمكن ان يقوم بها بنجاح . اما اذا كانت الضرورة تقتضى ان يسأل المفحوص فردياً ، وان تستعمل أسئلة تتبعية اذا كان السؤال الأول غير واضح . فان هذا يقتضى توفر قدر كبير من الخبرة والمهارة ، أى تدريباً خاصاً وطويلاً بما يتفق مع درجة حساسية الموقف الاختبارى ، ودقة القرار الذى يتخذ فى ضوء بيانات الاختبار .

وتشير بعض المؤسسات المسئولة عن طبع الاختبارات وتوزيعها ، وعلى الخصوص NFER فى إنجلترا ETS (*) فى أمريكا ان تتوافر فيمن يشتري الاختبارات شروط ومؤهلات معينة - قد تصل الى الدكتوراه فى علم النفس قبل السماح ببيعه او توزيعه ، ناهيك عن القيود الصارمة على التوزيع والنشر التى تفرضها هذه المؤسسات وغيرها على الاختبارات حتى لا تفقد وظائفها الأساسية فى التقويم النفسى . ولك ان تقارن هذا بما يشيع فى بلادنا من فوضى حيث تباع الاختبارات وتشتري ، بل وتصور ، دون رقيب ، بل يصل الأمر الى حد اعادة طبعها كاملة فى ملاحق رسائل الماجستير والدكتوراه وفى هذا كله خروج على الميثاق الأخلاقى السيكولوجى لاستخدام الاختبارات .

٢ - الفة الفاحص بالاختبار : من أهم الواجبات التى ينبغى ان يراعيها الفاحص ان يتيح الفرصة لكل مفحوص ليظهر قدرته ، وليحصل على نتائج يمكن مقارنتها بنتائج مفحوصين آخرين ، فعندما تقوم عملية انتقاء ما على نتائج اختبار من الاختبارات فانها لن تكون عملية عادلة ما لم يتم اختبار الأشخاص جميعاً بطريقة موحدة ومتشابهة تماماً .

ولهذا كان من الضرورى ان يعرف الفاحص الاختبار الذى سيستعمله ، بل ويألفه الى أبعد درجة ممكنة ، اذا كنا نهدف الى تحقيق عملية اختبار

(*) هما

(NFER) National Foundation for Educational Research

(ETS) Educational Testing services

على التوالى .

بـقيقة منصفة - ومما يساعده على هذه المعرفة والألفة أن يقرأ كراسية التعليمات التي تصحب الاختبار عادة بعناية وأن يلتزم بها فيها التزاما تاما .

٢ - **الالتزام بالاتجاه العلمى :** على الفاحص أن يلتزم اتجاها علميا محايدا اثناء عملية الاختبار ، ذلك لأن المختبر غالبا ما يكون مهتما بالمفحوصين ، وقد يميل الى أن يراهم يحسنون صنعا فى الاختبار الذى يعطيهم اياه ، فينزح الى أن يساعدهم بنية حسنة ، بطريق مباشر أو بطرق خفية ، وهذا أكثر حدوثا فى مواقف الاختبار الفردية . ولكن على المختبر أن يلاحظ هذا فى نفسه ، وأن يتجنبه . ولا شك أن مما يساعد على هذا أن يلتزم المختبر بالتعليمات المضاحية للاختبار ، كما أشرنا ، فإذا ذكرت التعليمات مثلا أن سؤالا ما يكرر مرة واحدة . فلا يكرره المختبر أكثر من مرة واحدة فى أى ظرف من الظروف .

وكذلك على المختبر أن يلاحظ أى تعبيرات للوجه أو كلمات التشجيع أو الاشارات التى يراد بها حيث همة المفحوص ، هذا كله يضعف من قيمة الاختبار ومن تقنيته . وعلى هذا فمن المهم بالنسبة للفاحص أن يجعل هيئته وسمته وتعبيرات وجهه وإشاراته غير كاشفة . على ألا يعطى هذا المفحوص احساسا بأن المختبر لا يكثرث به ، بل لابد أن يحرص المختبر على أن يوحى للمفحوص ويؤكد له فى صمت اهتمامه بكل ما يقول .

٤ - **المحافظة على العلاقة بين الفاحص والمفحوص :** يلتفتنا هذا الى أهمية المحافظة على العلاقة بين الفاحص والمفحوص فى موقف الاختبار . فإذا كان مطلوبا من المفحوص أن يظهر افضل ما عنده فى الاختبار ، أى انه يتبعى أن يظهر باستمرار زغبته فى التعاون مع المختبر فانه من الضروري أن تقوم علاقة انسانية تتسم بالأطمئنان والثقة والقبول المتبادل بين طرفى هذه العملية - الاجتماعية فى أساسها - أى علاقة المفحوص والمختبر .

ونصارع فنشير الى أنه ليست هناك قواعد جديدة ولا نصائح تتبع لاقامة ذلك النوع من العلاقة بين الفاحص والمفحوص ، ولكننا نستطيع أن نقول بصفة عامة أن الشخص الذى يالف الناس ويألفونه ، والذى يوحى بالثقة وبالطمأنينة فى نفوس الآخرين ، يستطيع أن يجد من الوسائل ما يساعده على إنشاء تلك العلاقة بسعائها التى حددناها .

وهناك علامات تدل على سوء العلاقة الاختبارية ، أو على أن الفاحص غير موفق في إقامة العلاقة المطلوبة بمن يختبرهم ، ونذكر بعضاً من هذه العلامات حتى يستطيع الفاحص الفطن أن يتعرف عليها ويدركها ثم يستجيب لها بما يؤدي إلى اختفائها أو الإقلال منها . من هذه العلامات عدم انتباه المفحوصين أثناء القاء التعليمات ، أو الانتهاء قبل أن يحين الموعد المحدد ، أو عدم الاستقرار ، أو التماس أخطاء في الاختبار . أو اظهار اليأس من امكان الأداء أو العمل في الاختبار .

٥ - مسئولية الفاحص الأولى اختيار الاختبار الملائم للمفحوص : من الممارسات السيئة والخطيرة أن يعطى الفاحص للمفحوص اختباراً للقدرة مثلاً لا يناسبه ثم يؤدي فيه أداء سيئاً لأسباب غير مرتبطة بالأداء الاختباري ذاته . فاختبار المتأهية مثلاً لا يجب أن يعطى لمفحوصين يعانون من نقص في التحكم في حركات اليد ، وبالمثل لا يجب أن يعطى اختبار للميلول المهنية يلائم طلاب الجامعات لطلاب مرحلة التعليم الأساسي ما لم يكن قد أجرى تقنين جديد له يؤكد صلاحيته للمستوى التعليمي الجديد ، ويشمل ذلك بالطبع التأكد من أن تلاميذ مرحلة التعليم الأساسي يفهمون تعليماته واستلته .

ومن المعلوم في علم القياس النفسي وفنونه التطبيقية أن حصول المفحوص على درجة مضللة في اختبار غير ملائم أسوأ من عدم الحصول على هذه الدرجة على الإطلاق . ومسئولية اختيار الاختبار الملائم للمفحوص تظل في النهاية مسئولية الفاحص . وتبقى مسئوليته هذه حتى ولو طلب منه تطبيق اختبار بذاته من سلطة خارجية فعليه أن يستخدم مهارته وكفاءته في الحكم على ملاءمة الاختبار ، ويصدق هذا على مساعدى الباحثين الذين يقومون بجمع البيانات . ولعلنا في هذا الصدد لا نبالغ إذا قلنا أن بحثاً أو دراسة أو فحصاً يعتمد على بيانات زائفة أو سيئة أو مضللة أسوأ للعلم وأضل للممارسة العملية من عدم اجراء بحوث على الإطلاق .

طريقة اعطاء الاختبار :

١ - ظروف عملية الاختبار : هناك ظروف مادية لا بد أن تتوفر في الموقف الاختباري لما لها من آثار عميقة على سلوك الأشخاص من الفاحصين وبالتالي (التقويم النفسي)

على نتائج اختباراتهم ، من هذه الظروف التهوية والاضاءة ولهما اثر واضح على أداء الأشخاص فى الاختبارات ، سواء كان هذا الأداء كتابة أو عملاً . وكذلك من الضرورى توفير أماكن جلوس تساعد على القيام بهذا الأداء ، من حيث الاتساع وإمكان وضع الأدوات على مناضد بحيث يسهل تناولها واستعمالها وهكذا .

أما ترتيب أماكن الجلوس فلا بد أن يراعى فيه أن يساعد المفحوصين على الاستماع الجيد والسليم للإرشادات والتوجيهات . وهنا نشير إلى أن الصالات الكبيرة الواسعة لا تكون مناسبة للاختبارات الجماعية إلا إذا توفر العدد الكافى من المشرفين على أن يوزعوا توزيعاً مناسباً . وحتى مع توفر العدد الكافى من المشرفين فلا تزال هناك صعوبة لا بد أن نتغلب عليها وهى أن المفحوصين فى صالة ضخمة قد يحجمون عن السؤال أو الاستفسار من المشرفين خوفاً ورهباً ، وعلى هذا ، ينبغى عند إجراء اختبار جماعى فى صالة كبيرة أن يشجع المفحوصون على أن يسألوا ويستفسروا عما يعن لهم أو يلتبس عليهم فيما يتصل بما بين أيديهم من اختبار .

أما إذا انتقلنا إلى ظروف المفحوص نفسه ، من الناحية الجسمانية والنفسية ، فيجب أن نعطيها قدراً كبيراً من الاهتمام . وأول ما يلاحظ هو اختبار الوقت المناسب له من الناحية الجسمانية والنفسية ، لأن إعطاء الاختبار فى ظروف جسمانية أو انفعالية غير مناسبة ، لا تجعل أداءه عينة صادقة وممثلة لسلوكه .

ومما يساعد على تحقيق هذه الظروف توفير قسط مناسب من الراحة قبيل إعطاء الاختبارات ، ثم إثناءها بصورة منظمة دون ارتباك أو فوضى . وكذلك مراعاة أن لا تعطى الاختبارات عقب فترة من المجهود الجسمى أو العفى الشاق أو عقب التعرض لحالة انفعالية حادة ، إلا إذا كان الغرض من القياس ذلك .

أما إذا اضطررنا ، لسبب ما ، أن نعطي اختبارات فى ظروف غير مناسبة فيجب أن نعتبر نتائجها أولية وأن ننظر إليها نظرة شاقدة وأن نأخذها بغير قليل من التحفظ والحذر .

٢ - إدارة الجماعة وهيكليتها : لا يعطى الاختبار إلا لأفراد نتوقع أن

يفهموا توجيهات وتعليمات معينة ، كما نتوقع أن يسلكوا وفق هذه التوجيهات والتعليمات التي يوجهها الفاحص إليهم ، ولذا كان أساس الاختبار الجماعي ، من ناحية الفاحص ، هو ضبط الجماعة وإدارتها بطريقة تمكنه شخصيا من أداء واجباته المتصلة باعطاء الاختبار بدقة وحكمة ، ثم تمكن الجماعة من اظهار أحسن أداء ممكن على هذا الاختبار .

ومما يحقق للفاحص النجاح في هذا أن يقف أمام الجماعة بحيث يرى كل أعضائها وبحيث يستطيع كل أعضائها أن يروه ويسمعوه ، ويلقى تعليماته بعد أن يستقر ويسكن الجميع في أماكنها ، ويلاحظ أن تكون التعليمات بسيطة واضحة وأن يلقيها الواحدة اثر الأخرى بهدوء وأناة بحيث يكون كل جزء وكل فقرة من التعليمات متميزة ليتجنب سوء الفهم والخلط من جانب المفحوصين . وإذا كان انتباه الجماعة له كاملا ، وهو ما ينبغي أن يحرص عليه ، فإنه لن يحتاج إلى تكرار شيء مما يقول .

ومن الممكن أن يلجأ الفاحص إلى التمثيل ليزيد التعليمات وضوحا ، مالم يكن هناك نص في دليل الاختبار أو تعليمات موجهة إلى الفاحص نفسه تنص على منع التمثيل ، ففي هذه الحالة عليه أن يتجنب التمثيل والتوضيح . ومن أمثلة التوضيح مثلا أنه إذا كانت التعليمات تشير إلى قلب الصفحة الأولى من الاختبار ، فلا بأس من أن يمسك المختبر الاختبار عاليا ثم يقلب الصفحة ، أو أن يشير إلى المكان الذي ينبغي أن تكتب فيه البيانات الخاصة بالمفحوص ، وهكذا .

ومن المهم أن يكون واضحا أن للمفحوصين الحق دائما في أن يستفسروا ويسألوا ، ولعله من الأفضل أن يتوقع المختبر الأسئلة ويحاول الإجابة عنها قبل أن تسأل .

هل يتحقق ضبط الجماعة المفحوصة ويتميز اجراء الاختبار فيها عن طريق اتباع أسلوب صارم جامد من الفاحص ؟ لا شك أن مثل هذا الأسلوب يجعل أمر ضبط الجماعة وإدارتها سهلا لا مشقة منه ولا عناء على الفاحص إلا أن مثل هذا الأسلوب يخلع على جو الاختبار طابعا غير انساني ، ومن المهم أن لا ننسى في الموقف الاختباري أو في أي موقف اجتماعي بل وانساني

بصفة عامة ، أنه إذا كان ضبط الجماعة ضرورة لتسيير أمورها وتحقيق مطالبها ، فإن من الضروري أيضا أن يحس أعضاء الجماعة بأن من يوجههم أو يشرف عليهم إنسان يهتم بهم ويحرص عليهم وليس شخصا رسميا يقيم بينهم وبين نفسه استقارا تجعلهم ينفرون منه مما يمتد أثره الى ما بين أيديهم من عمل يتطلب منهم أدائه . والموقف الاختباري على أى حال ، لا يتطلب الاقدرا من الضبط يحققه قدر يسير من الرسمية فى العلاقات بين المفحوصين والفاحص . وقد تزيء درجة الرسمية هذه أو تنقص حسب أغراض الاختبار وظروفه ، ففى بعض الاختبارات التحصيلية - وخاصة فى مواقف الامتحانات - قد يحتاج الأمر لدرجة كبيرة من الرسمية والا حدثت الفوضى وشاع الغش ، ولعل ما يحدث فى وقتنا الحاضر فى الامتحانات المدرسية من مشكلات نظام يؤيد هذا القول ، وهو على أية حال أمر يحتاج لدراسة مستقلة .

ومن المهم أن نشير الى أن الفاحص لابد أن يرجع الى فطنته وتقديره ويعتمد على حصافته فى كثير من الأحيان اثناء الموقف الاختباري ، فاذا نشأت ظروف اضطرارية اثناء الاختبار ، كأن يمرض مفحوص فجأة ، مثلا ، فهنا لابد أن يحتكم الفاحص الى تقديره الخاص ، مراعىا أن الهدف من أى اختبار هو الحصول على معلومات مفيدة ودقيقة عن الشخص وله ، وليس مجرد اجراء اختبار كأنه عمل روتيني والفراغ منه على أى حال وبأية صورة .

٣ - التعليمات للمفحوص : ان اعطاء التعليمات والارشادات أهم مسئولية يتحملها الفاحص ، ذلك لأنه من الضروري فى الاختبارات المقنة أن يعطى المختبر التعليمات كما وردت فى الاختبار دون زيادة أو نقص ، فعلى الرغم من أن الاختبارات تقيس الفروق الفردية فان طرقها ووسائلها تكون مرسومة بحيث تلغى التفاوت والاختلاف والتباين فى كل خاصية من خواص المفحوص فيما عدا تلك الخاصية التى يريد أن يقيسها الاختبار بالفعل . وعلى هذا فانه من الأهمية بمكان بالنسبة للمفحوص أن تتساوى جميع الظروف - قدر الامكان - بين المفحوصين فيما عدا ذلك العامل أو تلك الخاصية التى يريد الاختبار قياسها .

وعادة ما نسمع التعليمات نفسها بالأسئلة والاستفسارات ، ولكن عندما يجيب الفاحص على الأسئلة عليه ألا يتعرض لتوضيح شيء أكثر مما فى

التعليمات ، اذ عليه ان يتذكر دائما ان نتائج مفحوصية قد تقارن بنتائج مفحوصين آخرين في اماكن وأوقات أخرى .

وعلى أى حال فان الاختبار الجيد يأخذ في الاعتبار كما اشرنا أكبر قدر من الأسئلة المتوقعة ويرتب اجابات وحلول لها حتى تكون طريقة اعطائه مقننة ، وبدون ذلك فان هذه الأسئلة قد يفهمها البعض على وجه بينما يفهمها البعض الآخر على وجه آخر ، ومن الضروري دائما تحديد المطلوب من المفحوصين بصورة واضحة محددة الى أبعد حد مستطاع .

وفي بعض الأحيان تكون طريقة اتباع المفحوص التعليمات المقننة مقصودا بها أن تؤثر في درجته ، فاذا أعطيت هذه التعليمات بطرق مختلفة ترتب على ذلك اختلافا في الأداء لم يكن في الحسبان .

ومن المسائل الهامة التي تتصل بتقنين تعليمات الاختبار ما يتصل بالتخمين . ويظهر التخمين بصورة أقوى وأوضح في اختبارات القدرة واختبارات التحصيل حيث يسأل المفحوص : « اذا لم أعرف الاجابة فهل من الممكن أن ألجأ الى التخمين ؟ » وفي بعض الأحيان تتضمن توجيهات الاختبارات نصوصا تبين الاجابة على مثل هذا السؤال ، ولكن حتى مع وجود مثل هذه النصوص فان الأمر لا يخلو من ابهام يتطلب التوضيح . وانه مما يتعارض مع اصول اعطاء الاختبارات أن يتطوع الفاحص باعطاء نصح وتوصية من عنده لم تنص عليه توجيهات الاختبار .

ما هو الموقف السليم بالنسبة للتخمين في الاختبارات ؟ يمكن أن نلخص ما ينبغي أن نعمله في مواجهة التخمين فيما يلي .

(١) على الفاحص أن يؤكد ما تشير اليه تعليمات الاختبار ، فاذا كانت هذه التعليمات تشجع التخمين اكده وشجع عليه أما اذا كانت تنص على عدم الاعتماد على التخمين فلعليه أن يظهر هذا .

(ب .) أما اذا لم تكن في دليل الاختبار اشارة الى التخمين فعلى الفاحص أن ينصح المفحوصين بأن يستعملوا ذكاءهم وتقديرهم فيما يختص بما اذا كان من الممكن أن يخمّنوا ، عندما تواجههم أسئلة أو عناصر من الاختبار لا يكونون واثقين من صحة الاجابة عنها ، بحيث تكون تخميناتهم

مستنيرة ، فلا يختارون الا ما يعتقدون انه اقرب الى الصواب ، ولا تكون المسألة كلها في يد الصدفة والتخمين غير المعقول .

(ح) من الضروري ان يتعلم اولئك الذين سوف تطبق عليهم الاختبارات وخاصة من الطلاب ، أن الاسراف في التخمين في اختبارات القدرة واختبارات التحصيل ليس في صالحهم على الاطلاق .

(د) من المعروف انه قد وضعت معادلات لتصحيح اثر التخمين في الاختبارات ، وعلى الفاحص ان يلجأ الى استخدام المعادلة المناسبة عند تصحيح الاختبارات ليفسر الدرجة التي يحصل عليها الفحوص تفسيراً معقولاً .

استثارة الدوافع لأخذ الاختبار :

هناك حقيقة هامة على كل مشتغل بالقياس النفسى ان يعيها وان يعمل وفقاً لها اولا وهى ان الشخص الذى نختبره لابد ان يهتم بنتائج الاختبار الذى يعطى له وان تكون لهذه النتائج قيمة عنده . ففى القياس النفسى - الذى يختلف عن القياس المادى - نجد ان الفحوص يضع نفسه فى القياس ، ومالم يهتم شخصياً بهذه العملية فأننا لا نستطيع قياسه وان لم يتحقق هذا فأننا لا يمكن اختباره بدرجة نطمئن فيها على ما نحصل عليه من بيانات عنه . واذا كانت الدوافع لا تمثل جانباً له أهمية فى انواع القياس المادى المختلفة ، مثل الطول والوزن ، فانه لها أهميتها الكبرى فى القياس النفسى ، وحتى يكون القياس النفسى دقيقاً لابد ان يقبل الفحوص على الاختبار ويتحرك نحوه . ولهذا كانت استثارة دوافع الفحوص للاختبار مسألة جديرة باهتمام الفاحصين وعنايتهم .

(١) الحوافز التى ترفع الدرجة : الحوافز التى تساعد على الأداء فى الاختبارات النفسية متعددة ومتنوعة ، فمنها مثلاً الرغبة فيها ، ومنها ذلك الحماس الذى ينتج من اندماج الذات والاستغراق الذى يعبر عن الرغبة فى الاحتفاظ باحترام الذات واحترام الآخرين . ومن هذه الحوافز أيضاً مجرى حب العمل والاستغفال بحل عناصر اختبار ما ، فمثل هذا الحب والاهتمام

يستثير الشخص الى بذل مجهود أكبر فيه . ولا نستطيع أن نفعل حافزا يقوم بدور هام في بعض الأحيان في دفع المفحوصين وتحريكهم للاداء على اختبار ما ، ذلك هو الرغبة في اطاعة السلطة وتنفيذ التزامها كما تتمثل في الفاحص وما يوجهه الى المفحوصين من تعليمات وتوجيهات ، وهذه الرغبة كما هو واضح متصلة بحاجات اجتماعية عميقة منها مثلا الحاجة الى الاعتراف والتقدير ، وكذلك الحاجة الى الضبط والنظام .

وتظهر الحاجة الى استثارة حوافز المفحوص ورفع دوافعه الى الدرجة الملائمة في اختبارات القدرة واختبارات التحصيل بصفة خاصة ، وذلك لأن أهم ما يعنينا من هذين النوعين من الاختبارات هو الحصول على احسن اداء ممكن . عليهم من المفحوصين ، أو الحصول على أعلى انتاجية ممكنة فيهم اذا استعملنا تعبيراً . المسحة الاقتصادية .

والذي لا شك فيه أن مراعاة ما ذكرناه ، وما سناقش فيما يلي من شروط وظروف في الموقف الاختباري ، مما يساعد على تحقيق ذلك المستوى الأنسب من الدوافع لأخذ الاختبار والاستمرار فيه بجهد متوازن مناسب . وقد بذلت محاولات متعددة لاستعمال انواع بسيطة من المكافآت لرفع مستوى الأداء على اختبارات القدرة بنوع خاص ، ومن امثلة هذه المكافآت البسيطة الجوائز ، والنقود ، وأحاديث التشجيع ، وقد اشارت الأبحاث التي أجريت في هذا الصدد (١٨ ، ٢٢ ، ٢٥) أن مثل هذه المكافآت لم تحدث أي تحسن في درجات اختبارات القدرة عن الدرجات التي حصل عليها المفحوصون الذين أخذوا الاختبارات نفسها في الظروف العادية لا إعطاء الاختبارات . ومما يلاحظ أنه من الممكن أن تتحسن الدرجات على الاختبار عندما تحول دوافع الشخص من التركيز على مثل تلك المكافآت البسيطة الى الاهتمام بدرجات الاختبار التي يحصل عليها .

وحتى في الاختبارات التي تعطى لأغراض البحث العلمي وليس لاتخاذ قرار بشأن المفحوصين فإنه لا بد أن يعطى الفاحص أو القائم بالبحث للمفحوصين اسبابا لأخذ الاختبار لها قيمة شخصية لهم ، وبدون هذا فإن أداءهم أو مشاركتهم في الاختبار تكون عملا لا معنى له ، مما ينعكس على نوع الأداء وبالتالي على البيانات التي نقيم عليها البحث كله .

(٢) الدوافع التي تخفض الدرجة : أما الدوافع التي تؤدي الى خفض الدرجة على الاختبار فنستطيع أن نحدد منها ما هو متعمد مقصود ، وما هو متصل بالرغبة الزائدة في الأداء الممتاز والتفوق وما يصاحب هذا من قلق وتوتر .

وليس عجيبا أن يعتمد مفحوص الى أن يكون أداءه سيئا بحيث تكون درجته على الاختبار أقل كثيرا مما يستطيع ، ولكنه يفعل هذا لأنه ، مثلا ، لا يريد أن يحرم من عضوية جماعة يحبها ويحرص على أن يبقى فيها وحصوله على درجة عالية في الاختبار كفيلا بأن يبعده عن جماعته الى جماعة أخرى أو لأنه يتوقع أن تضاف اليه أعباء وواجبات إضافية نتيجة حصوله على درجة عالية في الاختبار وهكذا .

أما الدافع الشديد الى الامتياز والتفوق في الأداء على اختبار ما فقد يؤدي الى التأثير السيئ على هذا الأداء ، كما أن التوتر الزائد يؤدي الى أخطاء في الأداء كما نلاحظ في الاختبارات اللغوية أو الاختبارات التي تتطلب دقة حركية وتأزرا حسيا حركيا بنوع خاص .

ومن المهم أن نشير الى أن قلق الاختبار ، كما أظهرت الأبحاث التي أجريت عنه ، متعلم من معاملة الآباء والمدرسين للتلاميذ ، كما أنه متعلم من رفقاء التلاميذ وأصحابهم ويبدو أن قلق الاختبار موجود بدرجات تكاد تتساوى بين القادرين من التلاميذ وبين غير القادرين منهم كما أثبتت أبحاث ساراسون Sarason وزملائه (٢٩) ، مستعلمين في قياس قلق الاختبار اختبارا وضعوه لهذا الغرض خاصة .

وقد وجدت بعض التجارب التي أجريت على قلق الاختبار علاقة سالبة بين القلق والأداء في اختبارات التحصيل واختبارات الذكاء ، مما يؤيد الفرض الذي يذهب الى الإطلاق القلقين في موقف الاختبار يهبط مستوى أدائهم ، ولكن ، كما تشير استتازي (١١ : ٢٦ - ٢٨) لا ينبغي أن نأخذ هذه العلاقة السالبة بين القلق والأداء في الاختبار على أنها حتما إشارة الى علاقة سببية بينهما ، ذلك لأنه من المحتمل أن ينشأ عند المفحوص قلق لأن أدائه في الاختبار

كان ضعيفا ، وليس العكس . ومما تشير اليه نتائج الدراسات فى هذه الناحية ، ان درجة بساطة من القلق تساعد فى الأداء فى الاختبارات وفى التعلم ، بينما المستوى العالى من القلق يكون له اثر معطل . وتشير النتائج ايضا الى ان التعليمات الموجهة الى المفحوصين التى تتضمن توضيحا لاهمية الاختبار ونتائجه للمفحوص وتستدعى اندماجه فى الأداء فيه ، هذا النوع من التعليمات يحدث اثرا ايجابيا على أداء المفحوصين الذين يكون لديهم مستوى منخفض من القلق عندما يبدأون الأداء فى الاختبار ، بينما تحدث اثرا سينا عند أولئك المفحوصين الذين يبدأون بمستوى عال من القلق .

على ان العلاقة بين القلق وبين الأداء فى الاختبارات ، وكذلك العلاقة بين القلق وبين التعلم ، تعتبر من المسائل التى تحتاج الى بحوث وتجارب عديدة حتى تتحدد بصورة أدق وأوضح ، ويمكن الرجوع الى بعض نتائج البحوث فى هذا الصدد الى بعض المراجع المتخصصة (٧) .

ومن الطبيعى ان نتوقع ان كل موقف اختبارى يتضمن نوعا من التهديد، منه مثلا الخوف من الفشل ، أو الخوف من العقاب ، أو الخوف من الاتهام، أو من الاعتراف بقبول وجود سعة أو خاصية شخصية قد يكشف عنها الاختبار ، مثل العصائية أو الميل الى التسلط . لهذا ، كان من المهم أن يؤكد الفاحص للمفحوص ان النتائج سوف تستخدم لصالحه ، مع تأكيد الجوانب الايجابية بصفة خاصة من هذا الاستخدام .

التهيئة لأخذ الاختبار :

(١) ادراك المفحوص لغزى الاختبار : من المهم أن يرغب المفحوص فى أن تكون درجته فى الاختبار صادقة صحيحة ، ذلك لأن هذه الرغبة ضمان لأن تكون الدرجة صادقة فعلا ، ويساعد التعاون بين الفاحص والمفحوص على ايجاد هذه الرغبة ، ولهذا كان امرا يستحق أن يحرص المختبر على توفيره، وهو ليس عسير التحقق على أى حال . على أنه من المهم أن يكون واضحا ان التعاون الذى نقصده بين المختبر والمفحوص ليس معناه أن تحال المسئولية كلها وتنقل الى المفحوص ، وانما نقصد أن يجعل الفاحص المفحوص يحس

بإستمرار أنه مشارك له فى العمل مشاركة العضو العامل فى أى جماعة ،
ولذا كان من الضرورى أن نتذكر أن الموقف الاختبارى المثالى هو الموقف الذى
يكون المفحوص فيه شريكا فى اختبار نفسه ، كما سبقت الإشارة عند الكلام عن
ذوابع المفحوص .

وعلى الفاحص أن يعطى المفحوص كل ثقته فيما يتصل بهدف الاختبار ،
وأن يصور له الاختبار على أنه فرصة متاحة وأداة علمية ميسرة يعرف عن
طريقها شيئا عن نفسه ، تجعله أقدر على اتخاذ القرار المناسب له ، وعلى
فهم ما يتخذ من قرارات متصلة به سواء فى العلاج أو الارشاد النفسى أو عند
التقدم لكلية أو معهد أو وظيفة ما . وإذا كان الطبيب البارع يشرح لمريضه
نوع الدواء الذى يعينه له والنتائج التى يتوقعها من تناوله الدواء ، فإن المختبر
الممتاز يجب أيضا أن يجعل الهدف من أخذ الاختبار والنتائج التى تترتب عليه
واضحة بالنسبة للمفحوص ، ولا شك أن هذا الوضوح يجعل المفحوص أكثر
حرصا على أن يجعل الاختبار يعطى صورة دقيقة بقدر الامكان عما يراه
قياسه .

(٢) دور المعلومات التمهيدية : يزداد الاعتراف فى الوقت الحاضر
بأهمية المعلومات السابقة التى يعطيها الفاحص للمفحوصين فى تهيئتهم
واعدادهم للاختبار ، وفى الارشاد النفسى مثلا (١٩) لم يعد اعطاء الاختبارات
مسألة روتينية يقوم بها المفحوصون قبل أن تبدأ عملية الارشاد النفسى ، بل
أصبح من الضرورى أن تسبق اعطاء الاختبارات مقابلات وغالبا ما يفضل
مقابلتان أو ثلاثة ، تتحدد فيها المشكلة ويناقش المختبر المفحوصين فى انصب
الاختبارات للمشكلة والغرض الذى من أجله يسعى هؤلاء المفحوصون الى
الارشاد النفسى . هذا ، وإن لمثل هذه المقابلات أثرا لا يمكن اغفاله وهو أنها
تساعد على الاقلال من قلق الاختبار الذى تحدثنا عنه .

(٣) دور التدريب على الاختبار : ظهرت فى السنوات الأخيرة
برامج ومؤسسات لتدريب المفحوصين على أخذ الاختبارات النفسية ، كما
ظهرت فى الأسواق سلاسل من « الاختبارات المصغرة » يتدرب عليها
المفحوصون قبل أن يتعرضوا لبرامج الانتقاء المهنى أو التعليمى . وازدادت
هذه الظاهرة شيوعا فى المجتمعات التى تلعب فيها الاختبارات دورا حاسما
فى هذا المجال ، وخاصة الولايات المتحدة الأمريكية .

وقد شغلت مسألة التدريب على الاختبارات النفسية الباحثين في ميدان القياس النفسى منذ نشأته لمبكرة ، وأجريت تجارب هامة للتحقق من فعالية التدريب على الاختبارات ، ويمكن أن نلخص نتائج هذه البحوث فيما يلى (٢٠ : ٧٨ - ٨٠) .

(أ) تختلف برامج التدريب على الاختبارات فى طبيعتها ، وتعتمد فعاليتها على التهيؤ المبدئى لدى المفحوص وعلى دوافعه . وقد لوحظ أن مجرد نوافر الفرصة للمفحوص لأن يعطى الاختبار مرتين له بعض الأثر ، وخاصة إذا لم تكن لديه فى الأصل فكرة عن الاختبار . كما لوحظ أيضا أن مجرد تخصيص عدد من الساعات لاعطاء اشارات وتلميحات عن طريقة اخذ الاختبار والتدريب عليه يمكن أن تزيد درجة المفحوص ما بين ٨ ، ١٠ نقاط .

(ب) تزداد فعالية التدريب كلما تحول من مجرد مراجعة سريعة الى مقرر منظم . فعند تدريب بعض الطلاب على حل بعض المسائل الرياضية لمدة ٢١ ساعة من الدروس داخل الفصل و ٢١ ساعة أخرى من الواجبات المنزلية أدى ذلك الى زيادة الدرجة بمتوسط ١٦ نقطة فى أحد اختبارات الاستعداد الرياضى . وحدث ما يشبه ذلك فى اختبارات الاستعداد اللغوى ، ويزداد متوسط عدد نقاط الدرجة بزيادة ساعات التدريب . ومعنى ذلك أن التدريب على القدرات والمهارات اللازمة للاداء الاختبارى ، والتنبيه الى حيل الاجابة على الاختبارات النفسية له أثر مؤكد .

(ج) يصدق قانون الغلة المتناقصة على نتائج التدريب على الاختبارات النفسية . ومعنى ذلك أن زيادة نقاط درجات الاختبارات تزيد بزيادة ساعات التدريب حتى يصل عدد هذه الساعات الى نقطة معينة لا يترتب عليها أى زيادة مناسبة فى درجات الاختبارات . فإذا كانت زيادة الدرجة بمقدار ١٠ نقاط فى الاختبار اللفظى تتطلب ١٢ ساعة من التدريب ، وفى الاختبار الرياضى تتطلب ٨ ساعات ، فإن هذه الزيادة بمقدار ٤٠ نقطة تتطلب ١١٨٥ ساعة للاختبار اللفظى ، ١٠٧ ساعات للاختبار الرياضى . ومعنى ذلك أننا نجد أن العائد - عند نقطة معينة - لا يتكافأ مع الجهد المبذول .

(د) يؤدى التدريب على الاختبارات النفسية الى ظهور أنواع جديدة

من الظلم الاجتماعى ، والى الخلل فى تكافؤ الفرص ، فلا شك ان القادرين وحدهم من حيث وفرة الوقت والمال هم الذين يتلقون هذه البرامج التدريبية مما يعينهم على اجتياز عوائق الانتقاء والالتحاق بالمؤسسات التعليمية (الكليات الجامعية مثلا) والمهنية التى يرغبون فيها ، بينما لا تتاح هذه الفرص لغير القادرين .

عملية الاختبار باعتبارها علاقة اجتماعية :

لا ينبغى أن يتصور الفاحص نفسه انسانا لا يفعل ولا يتأثر ولا يحيد عن التزامه بما تنص عليه واجباته وأعماله الرسمية فى الموقف الاختبارى . وذلك لأن عملية الاختبار فى المجالات النفسية ليست عملية قياس « شئ » أو « موضوع » ليس بينه وبين الفاحص أية وشيجة ، باستعمال أداة فنية دقيقة الى حد بعيد ، كما هو الحال فى قياس النواحي المادية التى يقوم بها العالم الطبيعى أو المهندس والتى يكون فيها الفاحص محايدا الى اقصى حد بالنسبة للموضوع المقاس . أما والموضوع المقاس هو الانسان فان عملية الاختبار لا يمكن أن تخلو من علاقات اجتماعية ونفسية على درجة كبيرة من التعقد بين الفاحص وبين الشخص موضوع الاختبار .

(١) ادراك الذات وادراك الآخر عند الفاحص والمفحوص : المسألة فى العلاقة بين الفاحص والمفحوص اكبر وأعمق من مجرد أن تكون حرصا من جانب الفاحص على أن يشجع « المفحوص » وأن يكون « لطيفا » معه (١١) ، (٢١) كما هو شائع فى التوصيات التى توجه الى الفاحصين . وانما الأمر يتضمن شخصين ، فى حالة الاختبار الفردى مثلا ، يمثل أحدهما ، وهو المفحوص ، شخصا عنده تصورات عن نفسه وعن الآخرين ، واقعية أو غير واقعية ، وتصورات عن الفاحص نفسه وعن الموقف الاختبارى ، وعنده كذلك مخاوف وشكوك وتوقعات وآمال . وانه لحامل ذلك كله الى الموقف الاختبارى حيث يؤثر فيه من غير شك . أما الشخص الآخر فهو الفاحص ، وهو انسان عنده أيضا توقعات وآمال ومخاوف وشكوك وان اختلفت نوعا ودرجة عما عند المفحوص ، بل وعنده تحيزات وتفضيلات لا بد أن تنعكس وتؤثر فى كل مرحلة من مراحل الاختبار : اعطائه وتصحيحه وتفسيره .

ومن الطبيعى أنه عندما يأتى شخصان بهذه الحالة التى صورناها

موجزين الى موقف اختبار أو قياس أو تقويم فلا بد أن نتوقع نشأة علاقات اجتماعية ونفسية متشابكة تجعل مهمة المختبر أكثر دقة ورهافة .

وقد يلى نذكر أمثلة لهذا النوع من العلاقات تكفى لأن تلفت النظر الى إمكان تفرد هذه العلاقات والى غموضها . وتساعد الفاحص فى دور التدريب أو الممارسة الفعلية على أن ينمى عنده حساسية لهذه العلاقات فى المواقف الاختبارية ، بحيث يكتشفها بسرعة فيواجهها بصورة تجنبه خطأ أو تجاوزا يمكن تجنبه بالالتفات الى مثل ما سنشير اليه ، وتطبيقه فى مران طويل فى مواقف اختبارية تجريبية وحقيقية .

قد يحس شخص محول ليجرى عليه اختبار معين أو عدة اختبارات أن السلطة التى حولته تستبد به وتميل الى أن تسيطر عليه ، ويكون رد الفعل عنده هو التمرد على هذه السلطة ، وعلى الفاحص الذى يمثل « سلطة » أخرى بالنسبة اليه . ويظهر هذا التمرد المنقول من السلطة الأصلية الى سلطة الفاحص فى مقاومات ومظاهر معينة اثناء الموقف الاختبارى ، كأن يعطى مثلا معلومات يعتقد أنها مضللة عن نفسه ، أو يعطى معلومات خاطئة عمدا ، أو يظهر اهمالا وعدم اكتراث فى الاجابة أو الاداء ايا كانت صورته ، أو يبدى اعتراضات كثيرة ، وما الى هذا من مظاهر المقاومة والفشل فى العلاقة بين المختبر والمفحوص التى تكلمنا عنها من قبل . ولا بد أن يدرك الفاحص هذا ، خاصة فى المواقف الفردية ، ويعمل على الاقلال من هذه المقاومة ومعالجتها بالاعتراف بوجودها وتقبلها وبراازها للمفحوص ، وتأكيد الضرر الذى سيلحق به من اثر هذه المقاومة ، ومظاهرها المعطلة على النتائج التى يعطيها الاختبار . والتى هى فى النهاية لصالح المفحوص ، وبعد ذلك يمكن أن يعطى الاختبار .

بل ان الشخص الذى يأتى بنفسه طالبا اجراء اختبار أو عدة اختبارات عليه ، ربما لا يخلو هو أيضا من صعوبة على الرغم من توفر الدافع لأخذ الاختبار عنده ، ولو ظاهريا على الأقل . والأغلب أن نتوقع تعاونا كبيرا بين مثل هذا الشخص وبين الفاحص ، ولكن قد تتعارض دوافع هذا المفحوص مع أهداف الاختبار ، فإذا كان الدافع عنده هو أن يتعرف على بعض جوانب من شخصيته . فإنه فى الوقت نفسه يخشى أن يكشف الاختبار ، أو الفاحص عن نواحي قصور عنده يحرص هو بكل وسيلة أن يبقيها خارج نطاق وعيه .

وأكثر من هذا ، فإن مجرد سعى المرء الى أخذ اختبار ، أو الى المساعدة النفسية ، فيه أساس باستقلاله الشخصى واحساسه بقيمته وقدرته بحيث يزيد هذا عنده القلق ازاء الاختبار الفاحص جميعا .

وقد يكون الدافع الظاهر ، أو المعلن ، للسعى لأخذ اختبار غير الدافع الحقيقى وراءه ، فالطالب الذى يريد أن يأخذ اختبارات لأغراض تعليمية أو التوجيه المهنى قد يكتشف الفاحص المحنك أن وراء هذه الرغبة الظاهرة مشكلات نفسية واجتماعية هى التى حركت هذا الطالب الى طلب أخذ الاختبار وغالبا ما يكون هذا الاكتشاف بداية مرحلة من العلاج أو الارشاد النفسى قد يطلب بعده معلومات من اختبارات ، وأنه لقمين أن يحسن استخدام هذه المعلومات بعد أن تحل المشكلات التى كانت محاولات ابعادها واخفائها تستنفد جهده وطاقته ، والتى كانت تعوق خطاه عن السير فى الاتجاه المناسب الذى يرتضيه .

ومن الممكن أن يغلب على المفحوص مظهر الخضوع الكامل للفاحص ، ويبدى من الرغبة فى التعاون أكثر كثيرا مما يقتضيه الموقف عادة ، ومثل هذا إذ يسلم نفسه للفاحص ، إنما هو فى الواقع يعبر عن ميله الى التخلّى عن مسؤوليته عن نفسه بالقائها كاملة على الفاحص ، بل أن هذا الخضوع قد يخفى وراءه عدوانية أو مقاومة لابد أن يعالجها الفاحص أن كان ذلك فى مكنته أو يحيلها الى من يستطيع ويوفق فيها .

أما عن الفاحص نفسه فانه ، كما قلنا ، يحمل بدوره الى الموقف الاختبارى تصورات عن نفسه وعن المفحوص ورغبات وشكوكا مما أشرنا اليه . وهذه قد تتوافق مع تلك التى عند المفحوص بحيث تدعم كل منها الأخرى ، وقد تتعارضان وتتضادان . فإذا تصادف أن ذلك المفحوص ، الذى يلبس لباس الخضوع ويتخلّى عن مسؤوليته عن نفسه للفاحص ، وقع مع فاحص يميل الى السيطرة على الناس فإن هذا الفاحص ، فى أغلب الظن ، لن يستطيع أن يكتشف أن هذه السلبيه أو الخضوع ليس الا حائلا يقيمه المفحوص ليحول بين المختبر وبين الوصول الى أعماق لا يريد له أن يصل اليها . أما إذا تصادف وكان الفاحص من النوع الذى لا يرحب بأن تحال عليه تلك المسؤولية ، ربما يسبب . فى نفسه فانه لن يشعر بالأمان فى التعامل مع المفحوص ،

وقد يلجأ الى الافراط فى استخدام الاختبارات وما تسفر عنه من نتائج لأن كثرة البيانات تعطيه احساسا أكبر من الأمن مع مثل ذلك المفحوص .

وقد يكون عند الفاحص دوافع كامنة لا يتبينها أثناء قيامه بعملية الاختبار نفسها أو وراء اشتغاله بالاختبار بصفة عامة ، فقد يكون اختياره لهذا النوع من العمل اشباعا لحاجات معينة كأن يستشعر الكفاءة الاجتماعية من خلال التفاعل الاجتماعى مع الآخرين وراء امتار الاختبارات ورسمياتها لأنه يحس بعدم القدرة الاجتماعية فى مواقفها الطبيعية ، أو قد يكون كمن قاوم شكوكا أو يجيب عن تساؤلات عن نفسه بمقارنة نتائج مفحوصيه بنفسه ، أو قد يكون فى تعامله مع المفحوصين ساعيا وراء المسند العاطفى وذلك بأن يبدى رقة زائفة معهم ويكون عظيم التسامح والتساهل معهم ، مما قد يجمعه بفعل جوانب ضمنية فى نتائج الاختبار قد يكون لها مغزاها بالنسبة له وللمفحوص على السواء (كالنقص اللغوى) . على أن هذا ليس معناه أن يتجنب الفاحص أى تعبير انفعالى مع المفحوص ، بل على العكس كما سبق أن اكدنا أنه لا يصح أن يصطنع البرود أو الجمود أو الرسمية المفرطة ، وإنما المهم هنا هو أن يكون الفاحص واعيا بدوافعه وحاجاته هو حتى يستطيع أن يبعد آثارها كلما استطاع عن التأثير فى الموقف الاختبارى الذى هو مسئول عنه ، على أن يحرص على اظهار التدعيم العاطفى المناسب الذى ييسر إقامة علاقة سليمة بينه وبين المفحوص ويقلل من مقاومته ويساعده لى استدعاء افضل وأصدق ما عنده فى أدائه فى الاختبار .

(٢) دور اللغة فى تكوين العلاقة بين الفاحص والمفحوص : لعل القارئ يلاحظ أن معظم ما تضمنه موضوع ادراك كل من الفاحص والمفحوص لذاته وللآخر يدخل فى باب التواصل غير non-verbal communication بل ان بعضه يدخل فى مجال ما لا يمكن التلفظ به ، univertalized ونشير هنا الى الدور المباشر للغة فى تكوين العلاقة بين الفاحص والمفحوص .

ولعل من الأمور الواضحة بذاتها أنه اذا لم يفهم كل من الفاحص والمفحوص لغة الآخر فان الموقف الاختبارى كله يدخل فى باب العبث أو منطق اللامعقول ، ولا نبالغ اذا قلنا أنه يكون أدخل الى فن الملهاة ، وهذا كله بالطبع قد يكون مأساة بالنسبة للقرارات التى تعتمد على نتائج مثل هذه الاختبارات .

ويصدق هذا القول على الباحث الذي يحمل أدواته ويذهب بها الى احدى القبائل البدائية المنعزلة ثم يقدم لهم تعليمات وأسئلة بلغة لا يفهمها الا هو ، كما يصدق على الباحث الآخر الذي يختار عيناته من اعمار معينة (اطفال صغار مثلا) او مستويات اقتصادية اجتماعية معينة ثم يطبق عليهم اختبارات لا يكاد يفهم تعليماتها او أسئلتها سواء ، بل قد لا يفهمها هو لأنها مترجمة ترجمة شبه حرفية عن لغة أخرى . نقول هذا القول لأنه يوجد في المكتبة العربية في الوقت الحاضر طوفان من الاختبارات لم يبذل فيها أصحابها جهدا الا محض الترجمة - التي قد تكون ركيكة - الى اللغة العربية، ونكاد نقول ان بعض هذه الاختبارات لا يكاد يفهمها لغويا أصحابها انفسهم .

بالطبع توجد لغة الايماءات والاشارات التي نجح علماء النفس في استخدامها مع العاجزين لغويا ومع الذين لا يتحدثون اللغة القومية ومع الاطفال الذين لم يكتمل نموهم اللغوي بعد ، الا ان لهذه اللغة اصولها التي يجب اتقانها والتمكن منها ، والا ادخلتنا في فن العبث او منطق اللامعقول .

وقد اثبتت البحوث التي اجريت في الولايات المتحدة (٢٠ : ٦٦ - ٦٧) - وهي مجتمع متعدد اللغات - ان الاطفال الذين يتقنون لغة معينة - كاللغة الأسبانية - يجب اختبارهم بهذه اللغة دون اللغة الانجليزية ، الا اذا كان الهدف من الاختبار قياس مدى اتقانهم اللغة الانجليزية . وبالطبع فاننا في اختبار الاطفال يجب ان نستخدم لغة منتقاة من قاموسهم اللغوي، ومن الافضل بالطبع - في هذه الحالة وامثالها كاختبار الاميين - استخدام اللغة العامية سواء في اعطاء التعليمات او في صياغة أسئلة الاختبار . وفي هذا الصدد لا بد للمتخصص في مجال القياس النفسي ان يكون على صلة وثيقة بالباحثين في مجال علم اللغة والمتخصصين في اللهجات المحلية ، وان يوفر لنفسه معاجم هذه اللهجات - ان وجدت . ومن المستحسن في هذه الحالات ان يكون الفاحص الذي يطبق الاختبار على هذه العينات من المتحدثين الاصليين بهذه اللهجة . حتى لا ينشأ عن سوء الفهم او سوء الاستخدام عائقا في فهم المفحوصين للاختبار .

الا ان اللغة لا تعمل في جو اجتماعي فارغ او اجوف ، فلا بد في جميع الأحوال من الدفء العاطفي وجو الصداقة الذي يوفره الفاحص للمفحوص

فى موقف الاختبار . وقد أثبتت البحوث التى أجريت على تحليل محتوى المقابلات الشخصية أنه حين تجرى مقابلات رسمية مع أطفال من سن ٨ سنوات يعيشون فى حي الزنوج بمدينة نيويورك (حى هارلم) لم ينتج هؤلاء الأطفال - إلا فيما ندر - استجابات تزيد عن كلمة واحدة ، سواء أجريت هذه المقابلات داخل المدرسة أو داخل المنزل . ولكن حين يستخدم الأسلوب غير الرسمى فى المقابلة ، كان يتم فى حضور بعض أصدقاء الطفل والجلوس معهم حيث يجلسون - على الأرض مثلا ، ومشاركتهم فى شرب علب العصائر التى يشربونها ثم إجراء المقابلة فى صورة حديث عادى ، فإن النتائج تثبت فى هذه الحالة أن الطفل الذى يحكم على لفته من خلال المواقف الرسمية للمقابلة بأنها غير ناضجة يندفع منطلقا فى حديث طلق صريح وغير متحفظ .

ولعل القول الأخير يصدق أكثر ما يصدق على الاختبارات التى تتناول الجوانب الوجدانية والاجتماعية لشخصية المفحوص (الميل والاتجاهات مثلا) . ولعل أشهر الأمثلة على ذلك تطبيق مقاييس الاتجاهات التى تتناول المسائل الاجتماعية الخلافية الحساسة كالاتجاه نحو تحديد النسل فى مصر - ان اجابات المفحوصين على هذه المقاييس تأتى مباشرة وقصيرة وتتفق مع قواعد « المرغوبة الاجتماعية » ، اذا لجأ الفاحص الى الأسلوب الرسمى ، ولكن اذا تبسط مع المفحوصين ، كان يجلس مع الفلاحين على « المصطبة » ، ويشرب معهم قهوتهم أو شايتهم ، فإن هذا اللون من « الأحاديث العادية » يطلق الألسنة من عقالمها ، ويجعل الموقف الاختبارى أكثر طبيعية ، وحينئذ تودى اللغة دورها الطبيعى فى تكوين العلاقة الطيبة بين الفاحص والمفحوص .

كيف يشوه الفاحصون نتائج الاختبارات ؟

يسجل تاريخ العلم امثلة على خداع الذات الذى قد يقع فيه باحثون وملاحظون مدربون . وكثير من هذه الأخطاء يحقق لهؤلاء الفاحصين نوعا من الخدمة الذاتية ، قد لا يكون من بينها المكافآت المادية المباشرة لهم ، وإنما قد يكون الأهم عن ذلك اشباع حاجات الفاحص الى تدعيم آرائه ومعتقداته .

وبالطبع فإن تشويه نتائج الاختبارات اثناء عملية التقويم النفسى خطأ بالغ ، ان لم يكن خطرا عظيما . لنفرض أن أحد الباحثين التربويين صمم (التقويم النفسى)

طريقة جديدة فى تدريس الفهم اللغوى ويرغب - ولو لاشعوريا - فى ان يثبت كفاءتها وفعاليتها . انه قد يطبق الصورة (ا) من اختبار الفهم اللغوى قبلها ، والصورة (ب) بعديا ، وبالطبع فانه فى هذه الحالة قد يضمن الحصول على نتائج ايجابية لصالح طريقة التدريس المقترحة اذا ذكر للمفحوصين اثناء القياس القبلى ان درجات الاختبار ليست لها اهمية ثم قرا عليهم تعليمات الاختبار بسرعة ، وربما يسىء قراءة زمن ساعته فيختصر زمن العمل فى الاختبار . اما اذا استخدم الفاحص عكس ذلك فانه قد يرفع من درجات مفحوصي المجموعة الضابطة اثناء القياس القبلى ، او درجات مفحوصي المجموعة التجريبية اثناء القياس البعدي . وفى الحالتين يحصل على نتائج مختلفة قد يساء فهمها وتاويلها .

وبعض اخطاء الفاحصين قد تكون مقصودة - وهذه يمكن ان تكون من قبيل الجرائم العلمية التى يذكر تاريخ العلم قليلا منها وصل بعضها بالباحثين الى حد تزيف الحقائق وفبركة النتائج (٨) . اما « الأخطاء التى تحركها الدوافع » ، motivated error فقد تكون غير مقصودة . فإى اجراء مقنن لتطبيق الاختبار ، مهما بلغت دقته لا يمنع حدوث بعض الآثار غير المقصودة فى درجات الاختبار حين يتوقع المعلم مثلا من تلميذه « النجيب » ان يؤدي أداء جيدا فيه ، وحين يضغط تحيز الفاحص عليه ضغطا لا شعوريا فيظهر الرغبة فى ان يحرز المفحوص من جنسه او دينه او بيئته او طبقته الاجتماعية تفوقا فى الاختبار . ويعود هذا مرة أخرى الى حاجة المفحوص الى معرفة « ذاته » وفهم « شخصيته » من حيث ميوله واتجاهاته وقيمه وتفضيلاته وتحيزاته وآرائه وأفكاره . وكلما ازداد استبصار الفاحص بذاته كان أكثر موضوعية ، فى مواجهة مواقف الاختبار .

الا ان هذا الاجراء يحتاج الى جهد منظم وتدريب هائى ، ولعلنا نعيد تنظيم طرق تدريس مقررات الاختبارات النفسية والقياس النفسى والتقسيم التربوى والنفسى فى الجامعات بحيث تهمل مسألة تدريب الفاحصين بجوانبها المختلفة التى تناولناها فى هذا الفصل محور برامج هذه المقررات . ويمكننا ان نستفيد أيضا من بعض ما يستخدمه الباحثون فى مجالات أخرى - كالطب - حيث يستخدم أسلوب « الفحص الأعمى المزدوج » double-blind لتجنب الوقوع فى خداع الذات . ومن ذلك أن يعطى لنصف المفحوصين عقار

محايد لا يختلف على أى نحو ظاهرى عن الدواء الفعلى بحيث لا يستطيع الطبيب أو المريض أو مساعد العمل التمييز بين المجموعتين ، ولا يتم هذا التمييز الا اثناء تحليل النتائج تبعاً لنظام ترميزهما . وبالطبع فان هذا الاجراء لا يزال صعب التنفيذ فى مجال البحوث النفسية والتربوية . والبديل هو ما يقترحه بعض الثقة فى الميدان مثل كرونباك (٢١ : ٦٨) - من ضرورة أن تطبق الاختبارات فى حالة استخدام المجموعات التجريبية والضابطة بواسطة فاحصين من الخارج لا يعرفون مقدماً أنواع المجموعات . ثم قبل تصحيح الاختبارات وتفسير نتائجها أن تخطط جميع الأوراق - الاختبارات القبليّة والبعديّة للمجموعات التجريبية والضابطة - خلطاً عشوائياً ، مع الاحتفاظ بمفتاح للتعرف على هذه الأوراق فيما بعد ، لا يكون متاحاً للملاحظين أو المصححين .

وتوجد عوامل مباشرة تكمن وراء تشويه نتائج الاختبارات ، سواء كان مقصوداً أو غير مقصود . ومن ذلك ما يترتب على درجات الاختبارات من نفع أو خسارة يتعرض لهما الفاحص ، ومن ذلك المعلم الذى تعتمد ترقيته أو مكافاته على نتائج فصله فقد يؤثر فى أداء تلاميذه فى الاختبارات بالغش المباشر أو التلويح به ، ويذكر هورن Horne وجارتى Garty فى بحثهما عام ١٩٨١ أمثلة كثيرة لما يحدث فى هذا الصدد فى المدارس الأمريكية (٢١ : ٦٨) .

الفصل الرابع

ثبات الاختبارات

معنى ثبات الاختبار :

يقصد بمصطلح الثبات *reliability* فى علم القياس النفسى دقة الاختبار فى القياس أو الملاحظة وعدم تناقصه مع نفسه ، واتساقه واطراده فيما يزودنا به من معلومات عن سلوك المفحوص . ولا بد للباحث أن يحدد درجة الموثوقية فى بيانات الاختبار كما تتمثل فى اتساق القياسات أو الملاحظات التى يحصل عليها . ويمكن التحقق من هذه الموثوقية بصفة عامة بإحدى وسيلتين : إما بتكرار تطبيق الاختبار نفسه على نفس المفحوصين ، أو باستخدام اختبارات معادلة .

وبالطبع فإن من حقائق القياس النفسى أن درجات الاختبارات تختلف من مرقف لآخر ، وقد يرجع ذلك الى اختلاف درجة الانتباه عند المفحوص ومقدار جهده المبذول من وقت لآخر ، وذلك بالنسبة للفترات الزمنية القصيرة ، أما إذا طالت الفترة فإن التغيرات فى درجات الاختبارات قد ترجع الى أو التعلم أو التغيرات فى الجوانب الصحية أو فى شخصية المفحوص . فإذا استخدمت اختبارات مختلفة – ولكن متماثلة – فإن الاختلاف فى الدرجات قد يرجع الى مدى ألفة المفحوص بالأسئلة فى كل من الحالتين .

ألا أن أخطر العوامل التى قد تؤدى الى اختلاف الدرجات فى المواقف الاختبارية المختلفة بالنسبة للاختبار الواحد ولنفس المفحوصين ، هو عامل الصدفة *chance* ، وهو عامل يرجع الى الظروف العشوائية التى لا يتحكم فيها الفاحص مثل تخمين المفحوص للإجابة أو التذبذبات والتقلبات الوقتية فى انتباهه ، ويطلق على هذا العامل خطأ القياس *error of measurement* .

وفى عملية التقويم النفسى لا يمكن أن يعتمد قرار الفاحص على أخطاء المصادفة هذه ، وخاصة إذا كان القرار مصيريا مثل انتقاء الشخص لمهنة

أو دراسة معينة ، أو اعطائه شهادة أو اجازة بالكفاءة أو الخبرة ، أو انتهاء العلاج بالنسبة لمريض ، أو الحاق طفل متخلف عقليا باحدى المؤسسات .

ويستخدم مفهوم الثبات بمعناه الواسع ليدل على مدى اعتماد الفروق الفردية فى درجات الاختبار على اخطاء الصدفة المتضمنة فى القياس (والتي تدل على عدم الثبات) أو على الاختلافات الحقيقية الاصلية فى الصفة أو الخاصية السيكولوجية موضع الاهتمام (والتي تدل على الصدق كما سينبين فى الفصل التالى) . ويمكن صياغة المشكلة صياغة فنية على النحو الآتى :

إذا افترضنا ان الدرجات التى يحصل عليها المفحوصين فى اختبار معين تمثل التباين الكلى total variance يصبح السؤال ماهى النسبة فى هذا التباين الكلى التى يمكن اعتبارها من نوع تباين الخطأ error variance وتلك التى ترجع الى التباين الحقيقى ؟ true variance ان الاجابة على هذا السؤال ليست سهلة كما يبدو لأن ما نعتبره أحيانا من نوع تباين الخطأ فى موقف معين يمكن اعتباره تباينا حقيقيا فى موقف آخر . فمثلا إذا كان اهتمامنا متجها نحو دراسة التقلب الوجدانى فان التغيرات اليومية فى درجات اختبار يقيس « الانشراح والانقباض » تصبح مرتبطة بفرض الاختبار نفسه وتكون بالتالى جزءا من التباين الحقيقى فى الدرجات ، أما إذا كان الاختبار يهدف الى قياس الخصائص الأكثر استقرارا فى الشخصية كالذكاء أو الميل العلمى تصبح هذه التغيرات اليومية ذاتها من نوع تباين الخطأ .

ويمكن التعبير عن انواع التباين المشار اليها بالمعادلة الآتية :

$$\text{التباين الكلى} = \text{التباين الحقيقى} + \text{تباين الخطأ} .$$

ومن هذه المعادلة نستنتج أنه إذا زاد تباين الخطأ يقل التباين الحقيقى ويؤدى ذلك الى تذبذب الاختبار فيما يقيس وعدم اتساقه أو بعبارة أخرى عدم ثباته ، والعكس صحيح أى إذا زاد التباين الحقيقى يقل تباين الخطأ ويؤدى ذلك الى مزيد من استقرار الاختبار وثباته .

وعلى ذلك يمكننا أن نعرف تباين الخطأ بأنه يتعلق بالشروط والظروف التى لا تتصل من قريب أو بعيد بفرض الاختبار . ولذلك فحينما يحاول المدرس

أو المرشد النفسى أن يقنن الاختبار بأن يحافظ على توحيد الظروف الاختبارية وذلك بضبط مكان الاختبار وتعليماته وزمنه وعلاقة بالمختبرين وغيرها من العوامل التى تحدثنا عنها فى الفصل السابق فإنه بذلك يقلل من فرص حدوث تباين الخطأ ويجعل درجة الاختبار أكثر ثباتا واستقرارا . ومع ذلك فلا يمكن اعتبار الاختبار أداة ثابتة ثباتا كاملا يصل الى درجة ثبات مقاييس الظواهر الطبيعية (كالتر للاطوال والكيلو جرام للوزان) . ولذلك لابد ان يحدد صانعو الاختبارات مقدار ثباتها فى كراسة التعليمات بحيث يكون دليلا للاخصائين النفسيين حينما تطبق هذه الاختبارات فى ظروف معيارية مقننة .

والمشكلة الجوهرية فى تحديد ثبات الاجراءات الاختبارية بالنسبة لمجموعة معينة من الأفراد هى فى التمييز بين التباين الحقيقى والتباين الذى يرجع الى الخطأ ، وعلاقة ذلك بنوع الاستنتاج الذى يرغب الباحث النفسى فى التوصل اليه من درجات الاختبار . وبعدما يتم هذا التمييز بين نوعى التباين يبدأ الباحث فى تصميم سلسلة من العمليات الامبريقية (التجريبية) والاحصائية تفيد فى الوصول الى افضل تقدير لنسبة كل من نوعى التباين .

مصادر التباين :

١ - الخصائص المستمرة والعامة فى الفرد : فالتباين فى درجات ثورنديك أول من أعد تصنيفا شاملا لهذه المصادر عام ١٩٥١ فى فصل عظيم كتبه عن « الثبات » فى كتاب لنديكويست الشهير (٢٦) ، وقد عدل ستانلى J.C. Stanley هذا التصنيف فى فصل جديد بنفس العنوان فى الطبعة الثانية من هذا الكتاب والتى حررها روبرت ثورنديك (٣١) ، وتشمل هذه القائمة المعدلة ما يلى (٣١ : ٣٦٤) :

١ - الخصائص المستمرة والعامة فى الفرد : فالتباين فى درجات الأفراد ينشأ فى العادة نتيجة لما يظهره الأفراد المختلفون من مقادير مختلفة فى سمات أو خصائص معينة ذات طبيعة مستمرة وعامة ، ومن ذلك الاختلاف فى مستوى قدرة معينة لها دورها فى الأداء فى عدد كبير من الاختبارات ويؤدى هذا الاختلاف الى فروق فردية ظاهرة ومن أمثلة ذلك القدرة على الاستدلال أو السهولة العددية أو الفهم اللغوى . وعلى هذا توجد فى الأفراد

اختلافات بقدر ما يوجد من سمات أو قدرات ، وهي كثيرة العدد باللغة التنوع لا يتسع المقام لحصرها ، ويمكن الرجوع إليها في المراجع المتخصصة (٩) . ويمكن القول أن أى أداء اختباري يعتمد جزئيا على القدرات والسمات التي تتداخل بالفعل في عدد من الاختبارات النفسية .

ولعل القارئ يلاحظ أن هذا المصدر من مصادر التباين يمثل خصائص مستمرة ومستقرة لدى الفرد ، ويؤدي بالضرورة إلى فروق فردية مستقرة وحقيقية في الأداء الاختباري ، وهو نوع من التباين المنتظم على نحو معين الباحث على تناوله في أى إجراء يتخذه عند تقدير الثبات .

ويضيف ستاللى الى فئة الخصائص المستمرة والعاماً مصدرين آخرين للتباين عما : القدرة العامة على فهم التعليمات والخبرة بأخذ الاختبارات test-wiseness وهما عاملان يكادان يؤثران في كل درجة اختبارية سواء أراد الفاحص أو لم يرد . ويلعب عامل فهم التعليمات دورا هاما على وجه الخصوص حين يكون الاختبار جديدا ، وكانت تعليماته معقدة . أما العامل الثانى فيتمثل فيما يظهره المفحوص « الخبير بالاختبار » من خصائص تختلف عن المفحوص الأقل خبرة ، فالأول يكون عادة أقل انفعالية ، وأمهر في التخمين ، وأكثر وعيا بالمنبهات الثانوية والخارجية التي تفيد في الإجابة على أسئلة الاختبارات . ولعلك تلاحظ أن وجسود تباين ناتج عن هذين المصدرين يعد من الأمور غير المرغوب فيها في ضوء الغرض من الاختبار إلا إذا ارتبطا بقدرة عقلية أساسية . ويمكن التقليل من أثرهما بصفة عامة عن طريق تحسين تعليمات الاختبار من ناحية وتقديم قدر من التدريب على أسئلة الاختبار من ناحية أخرى .

وتمثل هذه الفئة من مصادر التباين تحديا حقيقيا لمؤلفى الاختبارات ، لأنها تتطلب التحكم ، إلا إذا كان وجودها مطلوبا ومرغوبا فيه وفي هذه الحالة تصبح من نوع التباين الحقيقي . ويتطلب هذا من الباحث النفسى أن يقدم لمفحوصيه تعليمات على أعلى درجة من الوضوح وبأقل قدر من المنبهات الثانوية .

٢ - الخصائص المستمرة والخاصة في الفرد : قد ينشأ التباين في الدرجات الاختبارية نتيجة لخصائص معينة في الأفراد تنقسم بالاستمرار من

ناحية الا انها تكون على درجة من الخصوصية للمجال النوعى الذى نختبره ، ومن أمثلة ذلك ما نجده فى اختبارات التهجى أو القسمة المطولة دون سواها من الاختبارات . فبالإضافة الى التباين الذى يميز الأداء العام فى التهجى أو الحساب يوجد نوع آخر من التباين يرتبط بصورة معينة من صور الاختبار أو طريقة معينة فى هذا الاختبار كما هو الحال مثلا فى سلوك التهجى حين يكون على المفحوص أن يتهجى كلمات معينة شفويا أو يكتبها صحيحة فى الاملاء أو يتعرف على الأخطاء فى بعض الكلمات المطبوعة . واختبار القسمة المطولة الذى يتطلب انتاج الاستجابة أو اختيار الاجابة من عدة بدائل (الاختيار من متعدد) . وقد ينشأ عن ذلك نوع من التباين بالطبع . ويوجد مصدر آخر للتباين ينتمى الى هذه الفئة هو طبيعة عينة الأسئلة أو المفردات التى يتألف منها الاختبار وينشأ عن الفروق الفردية فى تفاصيل المعارف والمهارات التى لدى الأفراد المختلفين . وعلى هذا لو أعد الباحث اختبارين من عيتين من الأسئلة اختيرنا مستقلتين من نفس الأصل الاحصائى للأسئلة أو المفردات فقد لا يحصل الفرد الواحد على نفس الدرجة فيهما بسبب التنوع بين الأفراد فى المهارات والمعلومات اللازمة للإجابة على أسئلة كل من الاختبارين . وفى هذه الحالة تنشأ الحاجة الى المقارنة بين الأفراد فى اختبارات متماثلة فيما تقيس وتختلف فى انواع المفردات أو طريقة العرض للتحكم فى هذا المصدر باعتباره ينتمى الى تباين الخطأ .

٣ - الخصائص العامة المؤقتة فى الفرد أو الموقف الاختبارى : ومن أمثلة هذه الخصائص الصحة الجسمية ومقدار النوم الذى حصل عليه الفرد فى الليلة السابقة على الاختبار ، والتعب ، وحالة الدافعية ، والتوتر الانفعالى وغيرها من العوامل الداخلية فى الفرد والتى تؤثر على أدائه فى الاختبار ، بالإضافة الى العوامل الخارجية فى الموقف الاختبارى كدرجة حرارة الجو ، والاضاءة والتهوية وغيرها . ومن المعلوم ان أداء الفرد فى الاختبارات المختلفة يتأثر بهذه العوامل بدرجات مختلفة ، الا ان جميع الاختبارات تتأثر بها على نحو أو آخر . كما تختلف هذه العوامل فى درجة استمرارها وعموميتها ، فبعضها قد يتغير من يوم لآخر ، وبعضها الآخر من ساعة لأخرى ، وبعضها الثالث من لحظة لأخرى . ويمكن لنا أن نتصور هذه العوامل على انها تميز الفرد فى جلسة اختبار معينة دون سواها . ويحتاج الباحث النفسى فى مثل الأحوال الى التحقق من اتساق الأداء فى الاختبار

سواء في لحظة معينة هي لحظة تطبيق الاختبار ، أو من فترة زمنية الى أخرى (بعد عدة أيام أو عدة أسابيع) كما أن على الباحث أن يحدد ما اذا كانت هذه الخصائص المؤقتة من المصادر المنتظمة الحقيقية للتباين أم هي من مصادر تباين الخطأ .

٤ - الخصائص الخاصة المؤقتة في الفرد : ويشمل هذه الفئة العوامل ذات النطاق الضيق سواء من حيث الزمن أو المدى اذا قورنت بعوامل الفئة السابقة ، وبعضها يرتبط بالاختبار ككل ، وبعضها يرتبط بأسئلته ، ومن أمثلة النوع الأول ما يتصل بصعوبة التعليمات في حالة الاختبارات الجديدة التي قد تنشأ عنها فروق فردية في فهم طبيعة المهام التي يتألف منها الاختبار ، وقد يؤدي ذلك الى تباين عشوائي ينتمي الى تباين الخطأ . وكذلك قد تتطلب الاجابة على الاختبار استراتيجيات مختلفة يستخدمها الأفراد المختلفون ، بعضها قد يكون اكفاً من البعض الآخر ، وحينئذ ينشأ مصدر جديد لتباين الخطأ . وبالإضافة الى ذلك فإن الأداء في بعض الاختبارات - وخاصة اختبارات المهارات الحركية - قد يكون عرضة للتحسن نتيجة للممارسة ، ويؤثر في ذلك مستوى الممارسة لدى الفرد عند بدء الاختبار . كما قد توجد مجموعة من العوامل المتنوعة التي يشملها مصطلح « التأهب لأداء الاختبار » ، مثل التركيز على السرعة أو الدقة في اختبارات القدرات الموقوتة ، أو الوعي ببعض المنبهات في حالة الاختبارات الإدراكية ، أو الحالة المزاجية لدى الفرد في حالة مقاييس الاتجاهات أو قوائم الميول . وهذه العوامل - على الرغم من أهميتها - يمكن تجاهلها في الاختبارات البسيطة والمقننة ، إلا أنها تلعب دوراً أكبر في الاختبارات الموقوتة ، وفي الاختبارات التي تتكون من أنواع جديدة من الأسئلة أو المفردات ، وفي المقاييس التي تتطلب مقداراً كبيراً من التفسير الاستبطاني ، وفي الأداء الاختباري الذي لم يتجاوز المرحلة التكوينية (أي لم يصل فيه المفحوص الى اجابة نهائية ، وإنما تكون اجابته على هيئة مسودة مبدئية) .

وتوجد مجموعة أخرى من العوامل تنتمي الى هذا الفئة وتخص المفردات أو الأسئلة ذاتها التي يتكون منها الاختبار ، أو بضمع لحظات عن الزمن الكلي للاختبار وتمثل في التذبذبات الوقتية في الذاكرة أو الانتباه ،

أو التعويق المؤقت للأفراد ، أو التنوع الدورى فى مقدار الجهد المبذول ، وهى جميعا تنتمى الى تباين الخطأ .

٥ - العوامل المنتظمة أو العشوائية التى تؤثر فى تطبيق الاختبار أو

تصحيحه : قد ينشأ التباين من مصادر أخرى غير المفحوص ذاته وتنشأ عن ظروف اعطاء الاختبار وتصحيحه ، ومن ذلك زمن الاختبار وتعليماته ومقدار الضوضاء وتششت الانتباه وفشل الاضاءة ونقص مواد الاختبار (كالأقلام) وغيرها ، وهى جميعا تختلف من شخص لآخر أو من مجموعة لأخرى وهذا النوع من الاختلاف ينشأ على وجه الخصوص حين يطبق الاختبار الواحد فاحصون مختلفون ، أو فى الاختبارات التى تحتاج الى توقيت دقيق أو تلك التى تتطلب تعليمات معقدة للغاية . كما قد ينشأ هذا المصدر للتباين عند تصحيح أو تقدير الدرجة فى الاختبار وخاصة حين يعتمد ذلك على حكم الفاحص أو الملاحظ أو المصحح ، ومن ذلك اعطاء الدرجة أو الحكم فى اختبارات المقال ، ومقاييس التقدير أو الأساليب الاسقاطية أو فى أى اختبار يتطلب قدرا من التفسير أو الحكم من جانب الفاحص - أننا فى هذه الحالة قد نجدنا ازاء نوع من التباين يرجع فى أصله الى اختلاف الفاحصين أو الملاحظين أو المصححين ، أو الى التغيرات التى تطرأ على الفاحص الواحد من وقت لآخر ، أو الى تناقض لا يخضع للتحكم أو التنبؤ داخل الفاحص أو المصحح نفسه . وهذه العوامل جميعا تنتمى الى تباين الخطأ .

٦ - العوامل التى ترجع الى الصدفة البحتة : يوجد نوع آخر من

التباين لا تستطيع أن تعود به الى عامل بذاته (أو عوامل بذاتها) ولذلك يمكن أن نعتبر عوامله من النوع العشوائى أو الذى يرجع الى الصدفة ، ومن ذلك تخمين المفحوص للإجابة فى الاختبار حيث يلعب « الحظ » دورا واضحا فى انتقاء الإجابة (والتى قد تكون صحيحة) ، كما قد يرجع الى عوامل أخرى كثيرة غامضة تلعب دورها فى تشتيت انتباه المفحوص فى لحظة معينة وهذه العوامل تنتمى الى تباين الخطأ بمعناه المباشر .

طرق حساب الثبات :

توجد بالطبع أنواع من ثبات الاختبار بعدد مصادر التباين التى أشرنا

اليها ، أى بعدد الظروف التى يمكن أن تؤثر فى درجات هذه الاختبارات ولا تكون كلها أو بعضها متعلقة بفرض الاختبار ، وتوجد طرق تستخدم فى حساب الثبات يشير كل منها الى المصادر التى يمكن أن يرجع اليها تباين الخطأ . وهذه الطرق ليست بدائل يحل بعضها محل بعض ، وانما لكل منها استخداما الخاص حسب مصدر تباين الخطأ الذى نهتم به .

١ - اعادة الاختبار :

أشرنا الى أن من المصادر الأساسية لتباين الخطأ ما يتمثل فى التغيرات العشوائية التى تنتاب الأفراد من جلسة اختبار الى أخرى . وقد تنتج هذه التقلبات عن ظروف تطرأ على الموقف الاختبارى ولا يمكن التحكم فيها مثل التغيرات الفجائية فى الجو أو حدوث ضوضاء مفاجئة (وغيرها من مشتتات الانتباه) . وقد تنتج عن تغيرات فى ظروف المفحوص نفسه كما تتمثل فى حالات المرض والتعب والاضطراب الانفعالى والقلق وغيرها . وهذه العوامل تدل على المدى الذى تتأثر به درجات الاختبار بالتغيرات اليومية العشوائية فى المفحوص أو فى ظروف الموقف الاختبارى .

ومن الواضح أن درجة الاستقرار الرقوى Temporal Stability تعتمد الى حد ما على الفترة الزمنية التى تنقضى على تطبيق الاختبار والتى نقدر بعدها ثبات الاختبار . وتوجد أمثلة عديدة لاختبارات لها ثبات مرتفع اذا قدر هذا الثبات بعد فترات زمنية قصيرة (أيام قلائل أو اسابيع قلائل) الا أن ثباتها ينخفض انخفاضاً شديداً عندما تمتد هذه الفترات الى عدة سنوات . وكثير من اختبارات ذكاء أطفال ما قبل المدرسة لها ثبات مرتفع خلال فترة ما قبل المدرسة ولكنها تفشل فى التنبؤ بذكاء الأطفال فى مرحلة الطفولة المتأخرة أو المراهقة أو النضج بسبب ما يطرأ على الطفل من تغيرات ارتقائية فى النواحي العقلية .

ولا شك فى أن التذبذبات ذات المدى الزمنى القصير يمكن اعتبارها من نوع تباين الخطأ فى درجة الاختبار . ولذلك ففى الممارسة العملية عندما نقوم بتقدير مثل هذا النوع من الثبات باعادة الاختبار لابد من أن تكون الفترة الزمنية بين مرتى التطبيق قصيرة ، وتكون أقصر عند اختبارنا للأطفال

الصغار منها عند اختبارنا للكبار بسبب سرعة معدل النمو في الطفولة عنه في مراحل النمو الأخرى . وبالنسبة لأي نوع من المفحوصين فإن قواعد الخبرة تؤكد أنه يجب ألا تزيد فترة ما بين الاختبارين عن ستة أشهر .

هذا التحفظ مهم للغاية ، لأننا نعلم جميعاً أن السلوك الانساني في حالة تغير مستمر ، فإذا كانت هذه التغيرات تحدث في الأداء الاختباري للمفحوصين بعد انتضاء فترات زمنية طويلة فإنها يمكن أن تعد ضرباً من النمو أو التدهور في السلوك لا مجرد تغيرات عشوائية ، بل قد تنتمي هذه التغيرات الى ميدان من السلوك أوسع نطاقاً مما يتناوله الأداء الاختباري ذاته ، فالمفحوص قد يغير أدائه في اختبارات الاستعداد المدرسي أو الفهم الميكانيكي أو الحكم الشر بعد عشر سنوات بسبب تعرضه لخبرات غير عادية خلال هذه الفترة . بعبارة أخرى يمكن للفرد أن يتغير مركزه النسبي في الاختبار بعد هذه الفترة صعوداً أو هبوطاً بالنسبة لأقرانه من نفس العمر بسبب ظروف خاصة تعود الى المنزل أو المدرسة أو البيئة الاجتماعية أو بسبب المرض المزمن أو الاضطراب الانفعالي المستمر .

ويجب أن نشير الى أن العمليات النفسية ذاتها قد تختلف فيما بينها في مدى ما تظهره من تذبذبات وتقلبات وقتية . فمثلاً قد يتأثر اتزان الحركات الدقيقة للأصابع بأي تغير ولو كان طفيفاً في ظروف المفحوص أو في ظروف الموقف الاختباري بينما نجد الاستدلال لا يسهل التأثير فيه بالمؤثرات العارضة . ولذلك إذا أردنا أن نحصل على تقدير دقيق لثبات اختبار اتزان الأصابع يجب إعادة تطبيقه مرات عدة في قيام متوالية ، بينما قد تكفي جلسة اختبار واحدة لقياس الاستدلال دون حاجة الى إعادة التطبيق . ومعنى ذلك أنه من الضروري فهم طبيعة السلوك الذي يقيسه الاختبار قبل اختبار أنسب الطرق لحساب ثباته .

ويقدر معامل الثبات الذي يتأثر بتباين الخطأ الناتج عن التقلبات الوقتية بطريقة إعادة الاختبار test-retest . ويطبق الاختبار في المرة الثانية ، بصورته ذاتها على نفس الأفراد ، وقد يتطلب ذلك إعادة التطبيق عدة مرات في مناسبات مختلفة ، ويحسب معامل الارتباط بين درجات المفحوصين في مرتي التطبيق أو متوسط معاملات الارتباط إذا أعيد تطبيق

الاختبار عدة مرات ، ولعلك تلاحظ أن درجات الأفراد قد تختلف في المرة (أو المرات) المختلفة ، وهذا لا يهم كثيرا إنما المهم أن أوضاعهم النسبية لا تتغير كثيرا ، ومعامل الارتباط في جوهره هو طريقة تحديد الوضع النسبي ويسمى المعامل الناتج معامل الاستقرار coefficient of stability .

٢٠

ورغم بساطة هذه الطريقة فإنها تتضمن بعض المشكلات . منها مثلا أن الممارسة قد تنتج قدرا من التحسن في درجات إعادة الاختبار . ومنها أيضا أنه لو كانت الفترة بين موقف الاختبار وأعادته قصيرة فإن المفحوص قد يستدعي بعض استجاباته السابقة ، أو بعبارة أخرى قد نحصل على نفس النمط من الإجابة من حيث الوضع النسبي للضوابط والخطأ ولكن ليس بسبب العملية النفسية موضع القياس ، وبالتالي لا تكون الدرجات التي نحصل عليها من موقف الاختبار مستقلة بعضها عن بعض ، كما أن معامل الارتباط بين مجموعتي النتائج قد يكون مرتفعا ارتفاعا زائفا . وكذلك من مشكلات طريقة إعادة الاختبار أن طبيعة الاختبار ذاتها تتغير بسبب إعادة وخاصة إذا كانت أسئلة الاختبار تتطلب استخدام بعض العمليات العليا كالاستدلال أو الابتكار . ففي مثل هذه الأسئلة نجد أن المفحوص إذا استطاع اكتشاف العلاقة أو أعطى استجابة جديدة أو توصل إلى حل المشكلة ، فإنه يمكن أن يستعيد إجابته ذاتها في المستقبل اعتمادا على الذاكرة دون التفكير ، ومعنى ذلك أن طبيعة الأداء في الموقفين تكون مختلفة . ومع ذلك فإن معامل الارتباط بين مرتى التطبيق قد يكون مرتفعا ارتفاعا زائفا ، ومع ذلك لا يدل على ثبات حقيقي . وعلى ذلك فإن هذه الطريقة لا تصلح إلا مع الاختبارات التي لا تتأثر تأثيرا كبيرا بالإعادة أو التكرار مثل اختبارات التمييز الحسي السمعى أو البصرى . أما بالنسبة لمعظم الاختبارات النفسية فإن طريقة إعادة الاختبار بالرغم من شيوعها في الاستخدام قد لا تكون ملائمة لها .

٢ - الصور المتكافئة أو البديلة :

من المؤلف لدى كثير من الطلاب أن يدخلوا الامتحان « فيسعدهم الخطء » بعدد كبير من الأسئلة يتناول الموضوعات التي أجمتوا بدراستها ومذاكرتها أكثر من غيرها ، وقد يحدث العكس في كثير من الأحيان هذا الموقف يدل على مصدر آخر من مصادر تباين الخطأ في درجات الاختبار وهو مدى اعتماد

درجة المفحوص فى الاختبار على عوامل تتعلق بطريقة اختبار الأسئلة . فإذا افترضنا أن قام باحث آخر بأعداد اختبار تتوافق فيه نفس مواصفات الاختبار الأول مع اختلاف فى المحتوى فالى أى حد تختلف درجة المفحوص فى الاختبارين ؟ الواقع أن الفرق فى الدرجة التى يحصل عليها نفس الأفراد فى هذين الاختبارين هو نوع الثبات التى يتأثر بعينة الأسئلة التى يتكون منها كل من بينهما . فمثلا إذا افترضنا أن باحثا قام بأعداد اختبار لقياس المعلومات الميكانيكية يتكون من ٥٠ سؤالاً ، ثم قام باحث آخر بأعداد اختبار آخر مكافئ له (أى يتكون من ٥٠ سؤالاً مختلفاً تقيس نفس العملية النهائية) ، وانعرض أن أسئلة كل من الاختبارين قد صيغت بنفس الدرجة من المهارة بحيث يتناول نفس المستوى من الصعوبة ، فإنه تبعا لعوامل الصدفة البحتة الناجمة عن الفروق الفردية فى الخبرات السابقة للمفحوصين قد تختلف الصعوبة النسبية لمجموعتى الأسئلة من شخص لآخر ، وعلى ذلك فإن الاختبار الأول قد يحتوى على عدد كبير من الأسئلة لم يالفه المفحوص (أ) بينما تكون الفته بأسئلة الاختبار الثانى أكبر . وقد يحدث العكس بالنسبة للمفحوص (ب) .

فإذا افترضنا أن هذين المفحوصين متساويان تقريبا فى معلوماتهم الميكانيكية (أى فى درجاتهم الحقيقية أو التباين الحقيقى) فإن المفحوص (ب) سوف يتفوق على المفحوص (أ) فى الاختبار الأول ، بينما سوف يتفوق (أ) على (ب) فى الاختبار الثانى . ومعنى ذلك أن الموقف النسبى للمفحوصين قد يتغير بل قد ينعكس ويمثل هذا نوعا من تباين الخطأ يرجع الى الصدفة الناتجة من طبيعة محتوى كل من صورتى الاختبار .

ويقدر معامل الثبات فى هذه الحالة بتطبيق صورتى الاختبار فى جلسة واحدة أو فى جلسيتين مختلفتين على نفس الأفراد ، ثم يحسب معامل الارتباط من الدرجات فى الاختبارين ويسمى معامل الثبات فى هذه الحالة بمعامل التكافؤ Coefficient of equivalence . والمهم فى جميع الحالات أن تكون صور الاختبار متكافئة فعلا ، أى تقيسان نفس العمليات النفسية بنفس المستوى من الصعوبة أو الألفة ، وأن تؤثر هذه العمليات النفسية فى الأداء فى الاختبار بنفس الطريقة . ولكن ، كما سبق أن أوضحنا ، قد تؤثر الصفات الخاصة تأثيرا مختلفا فى الاختبارين ومن ذلك المعلومات الخاصة التى قد

تساعد المفحوص فى الاجابة على سؤال معين من اسئلة احد الاختبارين ولكنها لا تساعده فى الاختبار الآخر .

وتباين الخطأ المتضمن هنا يدل على التغيرات فى الأداء من مجموعة معينة من الأسئلة الى مجموعة أخرى ، أى الخطأ الذى يرجع الى عينة الأسئلة ولكنه لا يتضمن تباين الخطأ الذى يعود الى العوامل الوقتية ، والذى سبق أن تحدثنا عنه ، والذى يتمثل فى التغيرات فى الأداء من موقف اختبارى الى آخر . ويمكن الجمع بين هذين النوعين من تباين الخطأ اذا قمنا بتطبيق الصورتين المتكافئتين من الاختبار فى مناسبتين مختلفتين (أى لا نطبقهما فى نفس الوقت) . ويمكن بهذه الطريقة تفادى بعض مشكلات طريقة اعادة الاختبار ذاتها . وفى هذه الحالة يدل معامل الثبات على كل من الاستقرار والتكافؤ معا ، ويسمى بهذا الاسم .

وعند اعداد الصور المكافئة equivalent او البديلة alternate من الاختبارات يجب أن نتأكد كما قلنا من أن هذه الصور فعلا متكافئة . والصور المكافئة هى فى جوهرها اختبارات تم اعداد كل منها على حدة وبطريقة مستقلة بحيث تتوافق فيها نفس المواصفات ، أى أن تحتوى على نفس العدد من الأسئلة . وأن تكون صياغة هذه الأسئلة متماثلة ، وأن تتضمن الأسئلة محتوى متشابها ، وأن تتعادل الأسئلة فى مستويات الصعوبة ، وأن تتفق الاختبارات فى جميع المظاهر الأخرى مثل التعليمات والزمن والأمثلة التوضيحية والشكل العام . ولا يمكن أن تعزى الفروق فى الدرجات من اختبار الى آخر الى تباين الخطأ الذى يرجع الى عينة محتوى الاختبار الا اذا كانت الصورتان متكافئتين او متعادلتين حقا .

ومن المؤكد أن حاجتنا الى الصور المتكافئة من الاختبارات تتعدى مجرد تحديد ثبات الاختبار . فالصور البديلة تفيد كثيرا فى الدراسات التى نجربها حول آثار عامل تجريبى معين فى الأداء الاختبارى وخاصة اذا تطلب الأمر قياسا قبليا وبعديا ، كما يفيد استخدام الصور البديلة فى اعطاء وسيلة لخفض احتمال التدريب أو الغش . الا أن أهم صعوبات هذه الطريقة تتمثل فى عدم امكان اعداد صور متكافئة لكثير من الاختبارات لأن ذلك يضاعف الجهد والوقت والتكلفة المبذولة فى بناء الاختبار . ومن المعروف أن اعداد

اختبار نفسى واحد يتطلب جهدا عظيما ، فما بالك باعداد أكثر من اختبار لقياس نفس السمة .

٣ - التجزئة النصفية :

يدل تجانس أسئلة الاختبار على مدى اتساق واطراد أداء المفحوص على جميع الأسئلة التى يتكون منها الاختبار . فإذا كان لدينا اختباران تتكون أسئلة الأول منهما من عمليات ضرب فقط بينما تتكون أسئلة الاختبار الثانى من العمليات الحسابية الأساسية الأربعة فإن الأداء فى الاختبار الأول سوف يكون أكثر اتساقا واطرادا بين أسئلته من الاختبار الثانى ، لأنه فى الاختبار الثانى (وهو اختبار غير متجانس الأسئلة) قد يحسن بعض المفحوصين الأداء فى عمليات الطرح مثلا بينما يسيئون أداء باقى العمليات . فإذا زاد عدم تجانس الأسئلة ، كأن يتكون الاختبار مثلا من ٤٠ سؤالاً يقيس كل عشرة منها عمليات نفسية مختلفة كالعلاقات المكانية والاستدلال الحسابى والمفردات والسرعة الإدراكية فإنه قد تكون العلاقة ضئيلة للغاية (ان كانت هناك علاقة) بين أداء المفحوص على هذه الأنماط الأربعة المختلفة من الأسئلة .

من ذلك نتبين ان درجة الشخص فى اختبار ما تكون أكثر وضوحا عندما نحصل عليها من اختبارات متجانسة البنية نسبيا . وتبرز مشكلة وضوح معنى الدرجة الاختبارية اذا عدنا مرة أخرى الى المثال الأخير ، فإذا حصل المفحوصان (١) و (ب) على درجة مقدارها ٢٠ فى هذا الاختبار فهل يمكننا ان نستنتج من ذلك أن أداءها متساو تماما فى هذا الاختبار ؟ الإجابة بالنفى دون شك . فقد ينجح المفحوص (١) مثلا فى الإجابة على أسئلة المفردات والسرعة الإدراكية ويفشل تماما فى الإجابة على أسئلة الاستدلال الحسابى والعلاقات المكانية . أما المفحوص (ب) فقد يحصل على درجته من الإجابة على ٥ أسئلة للسرعة الإدراكية و ٥ أسئلة أخرى للعلاقات المكانية و ١٠ أسئلة للاستدلال الحسابى ويفشل تماما فى الإجابة على أسئلة المفردات . ويوجد عدد كبير من الاحتمالات الأخرى للدرجات التى يحصل عليها المفحوصون من الاختبارات غير المتجانسة فى البنية . ومعنى ذلك أن درجة المفحوص فى مثل هذا الاختبار لها معان مختلفة ، بينما يكون معناها أكثر وضوحا فى الاختبارات المتجانسة .

(التقييم النفسى)

والدرجة في الاختبار المتجانس الأسئلة (معاني المفردات مثلا) قد لا يكون لها أيضا معنى واحد تماما في جميع الحالات ، وإنما قد تعنى مثلا أن المفحوص أجاب اجابة صحيحة على المفردات العشرين الأولى تقريبا اذا كانت هذه المفردات مرتبة حسب درجة الصعوبة ، وربما قد يترك كلمتين أو ثلاث من الكلمات السهلة أو قد يفشل في الاجابة عليها ، وقد ينجح في الاجابة على كلمة أو كلمتين أصعب وأبعد في الترتيب ومع ذلك فان هذه الاختلافات الفردية قليلة وهيئة اذا قورنت بتلك التي نلاحظها في الاختبارات الأقل تجانسا .

وهكذا يعد عدم تجانس أسئلة الاختبار مصدرا لتباين الخطأ ، وبالتالي يؤثر في ثبات الاختبار . ولكن هل يختلف تجانس الأسئلة كمصدر لتباين الخطأ عن المصدر الذي سبق الحديث عنه وهو ملائمة هيئة المحتوى عموما ؟ الواقع أن الفرق بين هذين المصدرين يمكن أن يتضح اذا اعطينا مثلا متطرفا ، لنفترض أن هناك اختبارا يقيس كل سؤال فيه عملية نفسية مختلفة ولا يرتبط اننى ارتباط بالأسئلة الأخرى التي يقيس كل منها عمليات أخرى . اننا في مثل هذه الحالة يمكن أن نجد صورة متكافئة من هذا الاختبار تحتوي على نفس انواع الأسئلة وتوزيعها على نفس النحو الذي هي عليه في الصورة الأولى . وعندما نعطي الصورتين المتكافئتين قد تتفق درجاتهما اتفاقا كاملا مما يدل من الوجهة النظرية على معامل ثبات مرتفع لا يتأثر تأثرا كبيرا بعينة الأسئلة كمصدر لتباين الخطأ . ولكن عندما نهتم بالثبات في ضوء تجانس الأسئلة نجد أن هذا المعامل يصل الى الصفر أو يقترب منه . حيث أن أطوار الاداء من سؤال الى آخر في أى من صورتى الاختبار قد لا يتجاوز حدود الصدفة .

وللتحقق من أثر عدم تجانس الأسئلة كمصدر لتباين الخطأ في الاختبار يلجأ الباحث الى تطبيق صورة واحدة من الاختبار مرة واحدة ، ثم يقوم عند تصحيح الاختبار بتجزئته الى نصفين وعندئذ يحصل على درجتين لكل مفحوص كل منهما في كل نصف من نصفي الاختبار يمكنه المقارنة بينهما ويسمى هذه الطريقة التجزئة النصفية split-half . ويحسب الباحث معامل الارتباط بين درجات الاختبار التي حصل عليها المفحوصون في نصفي الاختبار ، ويسمى معامل الثبات الناتج معامل الانساق الداخلى consistency

internal ، حيث أن الأمر لا يحتاج أكثر من تطبيق صورة واحدة للاختبار مرة واحدة فقط ، وعملية التجزئة تتم أثناء تصحيح الاختبار وليس أثناء تطبيقه .

ولعل أخطر المشكلات التي تواجهها طريقة التجزئة النصفية هي كيف نجزي الاختبار للحصول على أفضل قسمين للمقارنة ، أى قسمين متكافئين ، وخاصة إذا علمنا أن معظم الاختبارات لا يمكن تجزئتها إلى نصفين صالحين للمقارنة بينهما بمجرد تتبع التسلسل الرقمي للأسئلة ، وذلك بسبب الفروق في طبيعة الأسئلة ومستوى صعوبتها بالإضافة إلى الآثار التراكمية للحماس في الأداء أو العمل والممارسة والتعب والملل وغيرها من العوامل التي تختلف في بداية الاختبار عنها في نهايته . ولذلك فإن إحدى الطرق الدقيقة والموضوعية للحصول على قسمين متناظرين ومتكافئين في الاختبار أن يكون الاختبار في أعداده المبدئي قد تحددت فيه مستويات الصعوبة لكل سؤال وذلك بإيجاد النسب المثوية للأفراد الذين أجابوا عليه إجابة صحيحة ، وعندئذ توزع الأسئلة على قسمي الاختبار تبعاً لتكافؤ مستوى الصعوبة وتشابه المحتوى . فإذا كان ترتيب الأسئلة في الاختبار متسلسلاً تبعاً لمستويات السهولة أو الصعوبة فإن الباحث قد يستخدم طريقة صالحة لمعظم الأغراض وفي نفس الوقت لا تستغرق وقتاً طويلاً ولا تتطلب جهداً زائداً وهي تجزئة الاختبار إلى نصفين أحدهما يضم الأسئلة ذات الأرقام الزوجية والآخر يضم الأسئلة ذات الأرقام الفردية ، إلا أننا يجب أن نحذر أنه في حالات مجموعات الأسئلة التي تتعلق بمشكلة واحدة (مثل الأسئلة التي تتلوه فقرة معينة في اختبار لفهم قراءة) فإن هذه الفقرة بأسئلتها جميعاً يجب أن توضع في أحد قسمي الاختبار من الاختبار ، لأننا لو وزعنا مجموعة الأسئلة التي من هذا النوع على نصفي الاختبار فإن التشابه بينهما سوف يزداد زيادة مصطنعة أو زائفة ، وإى خطأ أو صواب في فهم الفقرة قد يؤثر في الإجابة على نصفي الاختبار ، ويؤدي ذلك إلى ارتفاع معامل الارتباط بين النصفين دون أن يدل في واقع الأمر على ثبات مرتفع .

لقد بينا أننا نحصل على معامل الاتساق الداخلي بحساب معامل الارتباط بين درجات نصفي الاختبار إلا أننا يجب أن نشير إلى أن مثل هذا

المعامل يعطينا ثباتا لنصف الاختبار فقط وليس للاختبار كله . وهذه المشكلة لا توجد بالطبع في الطريقتين السابقتين (إعادة الاختبار والصور المتكافئة) .

ونحن نعلم أنه إذا تساوت الظروف الأخرى فإن الاختبار الأطول في عدد أسئلته يكون أعلى ثباتا ، لأننا لو حصلنا على عينة أكبر من السلوك نصل إلى مقاييس أكثر دقة وأكثر استقرارا . ويمكن تقدير الأثر الذي تحدثه إطالة الاختبار أو اختصاره على معامل الثبات باستخدام المعادلة التنبؤية الشهيرة التي اقترحها سبيرمان وبراون Spearman-Brown وهي :

$$r = \frac{r_n}{1 + (n - 1)r}$$

حيث أن r = معامل الثبات الذي نريد الحصول عليه
 r_n = معامل الثبات الذي حصلنا عليه

n = عدد مرات إطالة الاختبار أو اختصاره أي أنه إذا زاد عدد الأسئلة من ٢٥ إلى ١٠٠ سؤال فإن n تكون ٤ أما إذا اختصر من ٦٠ إلى ٢٠ سؤال فإنها تصبح $\frac{1}{3}$.

ويمكن اختصار المعادلة حينما تطبق على معامل الثبات المقدر بطريقة التجزئة النصفية لأنها تتضمن إطالة الاختبار مرتين ، فتصبح على النحو الآتي !

$$r = \frac{r^2}{r + 1}$$

ألا إن معادلة سبيرمان - براون تفترض أن التغيرات في درجات نصفية الاختبار متساوية تماما أي تفترض تكافؤ ثباتهما ، وهو افتراض يصعب في كثير من الأحوال توافره حتى ولو بدا نصف الاختبار على درجة كبيرة من

التكافؤ بالفعل . وقد اقترح رولون Rulon تبسيط معادلة سبييرمان -
جراون مع وضع تباين نصفى الاختبار فى الاعتبار ، وصاغ لذلك المعادلة الآتية:

$$r = 1 - \left[\frac{E^2_f}{E^2_k} \right]$$

حيث E^2_f = تباين فروق درجات نصفى الاختبار .
 E^2_k = تباين الاختبار ككل .

وقد يستعاض عن ذلك باستخدام معادلة جتسمان Guttman
التي تستبعد هذا الافتراض وتصلح لحساب معامل الثبات مباشرة دون حاجة
الى حساب معامل الارتباط . ويمكن صياغة هذه المعادلة على النحو الآتى :

$$r = 2 \left[1 - \left(\frac{E^2_1 + E^2_2}{E^2_k} \right) \right]$$

وفى هذه المعادلة الرمز E^2_1 ، E^2_2 = تباين النصف (١) والنصف (٢)
فى الاختبار
 E^2_k = تباين الاختبار ككل .

٤ - طريقة كيودور - ريتشارد سون :

تصلح طريقة التجزئة النصفية لحساب معامل الاتساق الداخلى (الاتساق
بين نصفى الاختبار) ، أما بالنسبة لموضوع تجانس الأسئلة ذاتها والتي تؤلف
فى مجموعها بنية الاختبار فان بعض العلماء يرون أنه لا يدخل فى نطاق
مناقشة موضوع ثبات الاختبارات وانما هو من خصائص الاختبارات الجيدة
بصفة عامة الا أن هذا الراى لا يتفق مع التصور العام لمفهوم تباين الخطأ
الذى عرّفناه فما سبق ، فنحن ، دون شك ، فى حاجة الى معرفة مدى
استقرار استجابات المفحوص على مفردات الاختبار سؤالا بعد آخر وهذا

ما يسمى « الاتساق بين أسئلة الاختبار » ، interitem consistency ، وهو مصطلح آخر قد لا يعنى نفس ما يعنيه مصطلح معامل الاتساق الداخلى فى طريقة التجزئة النصفية ، وحتى لا يحدث الخلط بين الطريقتين تسمى مصدر التباين هنا بتجانس الاختبار homogeneity ، وهو مصدر هام لأن عدم تجانس أسئلة الاختبار وتنوعها الشديد قد يؤدي الى اختلاف أداء المفحوصين وتغيره فى المواقف المختلفة ، فقد يوجد لدينا اختباران لهما معاملات ثبات مرتفعة حسبت اما بالصور المتكافئة أو بالتجزئة النصفية ولكنهما قد يختلفان فى المعاملات الخاصة بدرجة تجانس الأسئلة فيهما ، وقد تظهر هذه المشكلة فى صورتى الاختبار أو نصفيه . . . والواقع أن الفرق بين معامل الاتساق الداخلى (الذى يحسب بالتجزئة النصفية كما سبق أن أوضحنا) وبين معامل التجانس الذى نشير اليه هنا يمكن اعتباره دليلا على مدى التجانس أو التباين فى محتوى الاختبار .

وأشهر الطرق لحساب معامل التجانس طريقة كيودر - ريتشارد سون وهذه الطريقة تتشابه مع طريقة التجزئة النصفية فى أنها تستخدم اجراء اختباريا واحدا لصورة واحدة من الاختبار ، ولكنها بدلا من أن تتطلب تجزئة الاختبار الى نصفين متكافئين فانها تعتمد على فحص أداء المفحوص على كل سؤال على حدة ، وأشهر المعادلات الاحصائية التى تستخدم فى هذا الصدد هى :

$$R = \frac{\sum_{i=1}^n \sum_{j=1}^n x_{ij}^2}{n \sum_{i=1}^n x_i^2}$$

حيث أن R = معامل ثبات الاختبار

n = عدد الأسئلة التى يتكون منها الاختبار .

$\sum x_i^2$ = تباين الاختبار ككل .

محصوخ = يحسب باعداد نسبة الأفراد الذين أصابوا فى اجابتهم على كل سؤال ونسبة الذين أخطأوا فيسه ثم نضرب

التسببتين في بعضهما ونجمع حواصل الضرب بالنسبة
لجميع الأسئلة .

ويؤكد كرونباك (٢٠) أن معامل التجانس كما يحسب بطريقة كيودر -
ريتشاردسون هو في الواقع متوسط معاملات الاتساق الداخلي الناتجة عن
تجزئات متعددة للاختبار . أما معامل الاتساق الداخلي فيعتمد على تجزئة
تهدف إلى الحصول فقط على مجموعتين متساويتين أو متكافئتين من الأسئلة،
والتي تعطى معاملًا للثبات بالرغم من عدم تجانس الاختبار ، ولعل من الأمثلة
المنطوقة على ذلك أن يكون كل سؤال زوجي وفردى في الاختبار يقيس عملية
نفسية مختلفة عن أي سؤال فردى وزوجي آخر في الاختبار .

ومن الواضح أن معادلة كيودر - ريتشاردسون تصلح للاختبارات التي
تتم الإجابة على أسئلتها بالثنائية ، من نوع الكل أو لا شيء مثل صواب -
خطأ ، نعم - لا ، موافق - معارض وغيرها ، ولهذا يقترح أحد مؤلفي هذا
الكتاب (٩) أن يحل الرمز (١) ، (ب) محل (ص) ، (خ) في المعادلة
الأصلية لتصبح أوسع في الاستخدام وتمتد إلى نطاق اختبارات الأداء المميز .

٥ - معامل ألفا لكرونباك :

كثير من الاختبارات النفسية لا يقوم على نظام الكل أو لا شيء المشار
إليه عند الحديث عن طريقة كيودر - ريتشاردسون . ففي التقرير الذاتي
قد يحصل المفحوص على درجة مختلفة تبعاً لدرجة إجابته على سؤال يطلب
تقدير مدى حدوث سلوك معين بأنه يحدث « عادة » أو « أحياناً » أو « قليلاً »
أو « نادراً » ، أو عند الإجابة على سؤال في مقياس للاتجاهات بقوله إزاء
عبارة معينة بأنه « موافق جداً » أو « موافق » أو « لا رأي له » أو « معارض »
أو « معارض جداً » . وبالطبع فإن هذا النوع من الاختبارات قد يكون أكثر
حاجة إلى تحديد معامل التجانس فيه أكثر من اختبارات الأداء الأقصى التي
هي في العادة من نوع الكل أو لا شيء . ولهذا اقترح كرونباك منذ عام
١٩٥١ معادلة تطورت بعد ذلك على يد كايزر وميشيل عام ١٩٧٥ أطلق عليها
معامل ألفا Coefficient Alpha ، وفيها يحل مجموع تباينات درجات الأسئلة
محل مج ص خ (أو مج ا ب إذا شئت) . وصيغت المعادلة كما يلي :

$$r = \frac{\left(\frac{n}{n-1} \right)}{\frac{E^2 - \text{مج} (E^2, 1, 2, \dots, n)}{E^2}}$$

حيث الرمز مج $E^2, 1, 2, 3, \dots, n$ = مجموع تباينات الاسئلة من السؤال الأول في الاختبار الى السؤال الأخير .

٦ - ثبات الفاحصين والمصححين :

بعد أن نستبعد الأنواع الثلاثة السابقة من تباين القضا من التباين الكلي فإن التباين المتبقى قد يرجع الى الفروق الفردية الحقيقية موضوع الاهتمام ، وبعضه الآخر الى عوامل أخرى غير مرتبطة بالاختبار ولكن يمكن ضبطها تجريبيا . فمثلا لا يسجل الباحثون أخطاء القياس الناتجة عن تطبيق الاختبار في ظروف مشتتة للانتباه . أو في زمن أطول أو أقصر مما يحدده كتيب الاختبار ، أو اعطاء تعليمات مختلفة عن التعليمات المحددة ، لأن هذه الأخطاء جميعا يمكن استبعادها والتغلب عليها في الموقف الاختباري نفسه اذا اتقن الفاحص فن اعطاء الاختبارات الذي عرضناه في الفصل الثالث . ولذلك فنحن لا نجد في العادة معاملات ثبات خاصة تشير الى « التباين الخاص بتشتت الانتباه » أو « تباين اختلاف الزمن » أو « تباين اختلاف التعليمات » الخ . فمعظم الاختبارات تتخذ اجراءات مقننة للتطبيق (أشرنا اليها في الفصل الثالث) والتصحيح (سنشير اليها في الفصل السادس) بحيث نفترض مع وجود هذه الاجراءات والالتزام بها أن « ثبات الفاحص » و « ثبات المصحح » يصل الى درجة كبيرة من الكفاءة تجعل الاختبار مهيئا في الأغراض التطبيقية والعملية وبالتالي لا نكون في حاجة الى تقدير مثل هذا النوع من الثبات الذي يرتبط في جوهره بموضوعية الملاحظة أو التصحيح . ويصدق هذا القول على وجه الخصوص بالنسبة للاختبارات الموضوعية الجماعية التي تطبقها على جماعات كبيرة والتي تصحدها الحاسبات الالكترونية . ففي هذه الحالات جميعا لا يتطلب الأمر الا التأكد من أن الناحص يتتبع اجراءات التطبيق والتصحيح تتبعاً دقيقاً وتكون المشكلة في جوهرها الضبط المحكم لظروف اجراء الاختبار .

الا اننا نلاحظ أنه في الاختبارات الفردية مثل ستانفورد - بينيه يلعب الممتحن دورا أكثر تعقيدا . لأن اجراءات الاختبارات في هذه الأحوال ليست مقننة تقنيا زامدا بل يعتمد معظمها على كفاءة تدريب الفاحص ، وعلى نجاحه في تكوين علاقة بينه وبين المفحوص ، وعلى دوافع المفحوص . وقد يتطلب الأمر أن يقوم الفاحص بتقويم أداء المفحوص خلال عملية تطبيق الاختبار لأن مثل هذا الأداء يحدد كيف يستمر الفاحص في موقف الاختبار . ولذلك فقد يختلف الفاحصون في النتائج التي يحصلون عليها من نفس المفحوصين حتى ولو توافرت في هؤلاء الفاحصين أفضل الشروط . وهذا الاختلاف قد يكون من نوع تباين الخطأ الذي يمكن أن ترجعه الى الفروق الفردية بين الفاحصين ويؤثر بالطبع في التباين الكلي لدرجات المفحوصين . ولذلك فاذا كانت الاختبارات التي نستخدمها من النوع الذي تلعب فيه فردية الفاحص دورا هاما ، او اذا كانت من النوع الذي يتأثر بالعوامل الذاتية وطرق الملاحظة المختلفة ، نكون في حاجة ماسة الى تقدير « ثبات الفاحص » وخاصة اذا كان من الضروري معرفة العلاقة بين تقديرات أكثر من فاحص لنفس المفحوص او المفحوصين .

وبالمثل نجد أن بعض الاختبارات تثير مشكلة « ثبات المصحح » وخاصة اختبارات المقال والاختبارات الاسقاطية بوجه عام ، وهي الأدوات التي تسمح بمقدار كبير من التفسير الذاتي من جانب المصحح وهو الفاحص أيضا في كثير من الأحوال .

ومن الواضح أن معامل ثبات المصحح يمكن الحصول عليه بحساب معامل الارتباط بين الدرجات التي يعطيها مصححان أو أكثر لنفس الأفراد في نفس الاختبارات . ومعنى ذلك أن كل مفحوص يحصل على درجتين أو أكثر من تصحيح اختبار واحد . وهذه الطريقة لا تتضمن بالطبع نواحي القصور الموجودة في طريقة اعادة الاختبار . اما معامل ثبات الفاحص فهو أكثر صعوبة لأنه بطبيعته يتضمن اعادة اختبار من جانب أكثر من فاحص ، وبالتالي فإن قيام أكثر من فاحص باعادة الاختبار يعنى الوقوع في جوانب القصور التي سبق الحديث عنها في طريقة اعادة الاختبار التي يقوم بها نفس الفاحصين . بالاضافة الى أن معامل الارتباط الناتج يعكس عوامل الاستقرار الوقتي وثبات المصحح معا . ويمكن اللجوء الى نمط مختلف من التخطيط

التجريبى يعتمد على تحليل التباين لعزل تباين الخطأ الذى يمكن ارجاعه الى الفاحصين . وعلى كل فيمكن ان نعتبر كلا من معامل ثبات المصحح ومعامل ثبات الفاحص من نوع واحد ونطلق عليه اسم معامل الموضوعية . Coefficient of objectivity

ويمكن ان نلخص أنواع معاملات الثبات وطرق تقديرها فى الجدول رقم (١) .

جدول (١) أنواع تباين الخطأ وعلاقته بطرق حساب الثبات (عن ١١ : ١١٨) .

المعامل	طريقة تقديره	تباين الخطأ
معامل الاستقرار	اعادة الاختبار بنفس صورته على نفس المفحوصين فى مناسبة مختلفة	عينة الوقت
معامل التكافؤ	تطبيق الاختبار مع صورة مكافئة له فى نفس الوقت	عينة المحتوى
معامل الاستقرار والتكافؤ معا	تطبيق الاختبار ثم تطبيق صورة مكافئة له فى مناسبة مختلفة	عينة الوقت ، عينة المحتوى
معامل الاتساق الداخلى	التجزئة النصفية	عينة المحتوى ، تجانس المحتوى
معامل التجانس	تحليل المفردات بطريقة كيودر - ريتشاردسون أو بطريقة الفا	عينة المحتوى تجانس المحتوى
معامل الموضوعية	تكرار المصحح أو الفاحص ذاتية المصحح والفاحص عينة الوقت	ذاتية الفاحص عينة الوقت

العوامل المؤثرة فى ثبات الاختبار :

١ - درجة اعتماد الاختبار على الزمن والسرعة :

تمثل الاختبارات ذات الزمن المحدد مشكلة من نوع خاص فى موضوع

الثبات ، وخاصة اذا كانت الفروق الفردية فى درجات هذه الاختبارات تعتمد على سرعة الأداء . ولذلك لا يمكن استخدام طريقتى التجزئة النصفية أو كيودر - ريتشاردسون لحساب ثبات هذه الاختبارات لأن هاتين الطريقتين تعتمدان على الاطراد فى ارتكاب الأخطاء . فمثلا اذا كان الاختبار يتكون من ٥٠ سؤالا من النوع الذى يعتمد على السرعة اعتمادا كليا (أى على عدد الأسئلة التى يحاولها المفحوص فى جلسة الاختبار) كاختبار النقر أو اختبار العمليات العددية لطلاب الثانوية العامة مثلا فاننا نجد ان المفحوص الذى يحصل على الدرجة ٣٦ مثلا تكون درجته فى كل من النصف الفردى والنصف الزوجى من الاختبار ١٨ ، بعبارة أخرى تتساوى تقريبا درجته فى نصفى الاختبار الا اذا تأثرت هذه الدرجة بالأخطاء العارضة التى ترجع الى اهماله لعدد قليل من هذه الأسئلة . وفيما عدا ذلك فان الارتباط بين الدرجات الفردية والدرجات الزوجية يكون اقرب الى الارتباط التام ، وهو معامل زائف لا يدل بحال من الأحوال على ثبات مرتفع الاختبار .

ولذلك فافضل طرق حساب الثبات لهذه الاختبارات هى طريقتا اعادة الاختبار واستخدام الصور المتكافئة . ويمكن اللجوء الى طريقة التجزئة النصفية بشرط ان تتم التجزئة على اساس السرعة فى الأداء لا على أساس عدد الأخطاء . بعبارة أخرى يجب ان تعتمد درجات نصفى الاختبار على مجموعات من الأسئلة ذات زمن محدد منفصل بعضها عن بعض . ويمكن تحقيق ذلك باستخدام جزئين متكافئين من الاختبار يعطى كل جزء منهما فى نصف الزمن الكلى للاختبار . وهذه الطريقة اقرب الى طريقة تطبيق الصور المتكافئة من الاختبار فى نفس الجلسة أو بعدها مباشرة ، الا ان كل صورة هى نصف اختبار فى حين ان الدرجة التى يحصل عليها المفحوص هى درجته فى الاختبار الكلى (أى فى نصفيه) ، ولذلك لابد من استخدام معادلة سبيرمان - براون أو غيرها من المعادلات المناسبة لايجاد ثبات الاختبار ككل .

وإذا استحال تطبيق نصفين للاختبار مستقلين فيمكن تقسيم الزمن الكلى الى أكثر من نصفين ثم إعطاء درجة لكل قسم . ويمكن ان يتم ذلك بأن نطلب من المفحوص ان يضع علامة معينة امام السؤال الذى يكون قد وقف عنده عندما يعطى المتن إشارة معينة متفق عليها مقدما . ويمكن ان نجعل عدد المفردات التى اجاب عليها المفحوص اجابة صحيحة فى القسمين الأول

والرابع مثلاً وتعتبرها مجموعة مستقلة ثم نجمع اجابات القسمين الثانى والثالث فى مجموعة أخرى بحيث يمثل كل من هاتين المجموعتين نصف اختبار . ومثل هذا النوع من التقسيم الرباعى يحقق نوعاً من التوازن بين الآثار التراكمية للممارسة والتعب وغيرها من العوامل . وتصلح هذه الطريقة فى الاختبارات الأخرى التى لا تعتمد على السرعة وخاصة حينما لا تكون الأسئلة مرتبة حسب مستويات الصعوبة .

والسؤال هو : متى يمكن أن نعتبر الاختبار من النوع الموقوت وبالتالى يجب أن نأخذ فى اعتبارنا ما سبق مناقشته ؟ من الواضح أن مجرد استخدام زمن محدد للاختبار لا يعنى أنه اختبار سرعة لأنه إذا انتهى جميع المفحوصين من الاجابة فى الزمن الذى يحدده الاختبار فإن السرعة لا تلعب دوراً فى تحديد الدرجة . وحتى إذا أتم جميع المفحوصين ٤٠ سؤالاً من ٥٠ فإن الفروق الفردية التى ترجع الى السرعة تصبح لا وجود لها رغم أنه لا يوجد مفحوص واحد أجاب على جميع أسئلة الاختبار . ولذلك لا يمكن أن نعتبر مجرد محاولة الاختبار كله أو عدم محاولته دليلاً على السرعة فى مقابل القوة ، وإنما لابد من اللجوء الى وسائل أكثر دقة لتحديد المدى الذى يمكن أن نرجع فيه الفروق الفردية فى الأداء الاختبارى الى السرعة ويتم ذلك بتحديد النسبة التى يمكن أن نرجعها الى التباين فى السرعة من التباين الكلى لدرجات الاختبار . ويمكن تحديد هذه النسبة بسهولة بإيجاد تباين درجات الأسئلة التى أجاب عليها جميع المفحوصين وقسمة هذا المقدار على تباين الدرجات الكلية للاختبار . ويمكن صياغة ذلك فى صورة المعادلة الآتية :

تباين الدرجات التى تمت الاجابة عليها (ع ٢ ص)

(ع ٢ ص) تباين السرعة =

التباين الكلى لدرجات الاختبار (ع ٢ ك)

فإذا حدث وأجاب جميع المفحوصين على ٤٠ سؤالاً أصبح بسط المعادلة صفراً حيث لا توجد فروق فردية (أو تباين) فى عدد الأسئلة التى تمت الاجابة عليها (ع ٢ ص = صفر) ، وبالتالى يصبح ناتج المعادلة مساوياً للصفر ولا يتحقق ذلك بالطبع فى اختبارات القوة والتى لا تتأثر مطلقاً بالسرعة . ومن ناحية أخرى إذا كان التباين الكلى للاختبار (ع ٢ ك) يرجع

الى فروق فردية فى السرعة فقط ولا يتأثر مطلقا بالقسوة فان تباين الدرجات التى قد تكون تمت الاجابة عليها (ع^٢ ص) يصبح مساويا للتباين الكلى ويكون ناتج المعادلة مساويا للواحد الصحيح أى يكون الاختبار اختبار سرعة بحث .

وهناك طرق احصائية أخرى عديدة وأكثر دقة لتحديد هذه النسبة ولكن مناقشتها تفصيلا يخرج عن نطاق هذا الكتاب .

ويمكن ان نضرب مثلا بأثر السرعة على معاملات التناسق الداخلى من بحث قامت به أنستازى ودريك (فى ١١ : ١٥٢ - ١٢٣) على اختبارات القدرات العقلية الأولية لأعمار تتراوح بين ١١ سنة و ١٧ سنة . وفى هذا البحث حسبت معاملات الثبات بطريقة التجزئة النصفية أولا ثم عن طريق حساب معاملات ارتباط الدرجات فى نصفين منفصلين تبعا للسرعة والزمن وتتلخص النتائج فى الجدول رقم (٢) :

جدول رقم (٢) معاملات الثبات لبعض اختبارات بطارية القدرات العقلية الأولية

طريقة حساب معامل الثبات	معانى الكلمات	الاستدلال	المكان	العدد
التجزئة النصفية فى محاولة واحدة	٠.٩٤	٠.٩٦	٠.٩٠	٠.٩٢
استخدام نصفين مستقلين من حيث الزمن	٠.٩٠	٠.٨٧	٠.٧٥	٠.٨٢

ومن هذا الجدول يتضح أن اختبار معانى الكلمات هو فى أساسه اختبار قوة ، بينما يعتمد اختبار الاستدلال على السرعة الى حد ما ، أما اختبار المكان فيعتمد اعتمادا كبيرا على السرعة . كما يتضح أيضا أننا عندما نحسب معاملات الثبات بطريقة جيدة فان هذه المعاملات تنخفض فى اختبار المكان من ٠.٩٠ الى ٠.٧٥ ، وكذلك الشأن فى اختبار الاستدلال والعدد فان معاملات ثباتها تنخفض نسبيا ، أما اختبار معانى الكلمات الذى لا يعتمد على السرعة فان معامل ثباته لم يختلف اختلافا له دلالة عندما حسب بكل من الطريقتين أنفتى الذكر .

أما إذا حسب معامل ثبات الاختبارات الموقوتة بطريقة التجزئة النصفية،
نم صحح أثر التجزئة بمعادلة سبيرمان - براون فان جنكسون Guttman
يقترح تصحيحا اضافيا بالمعادلة الآتية :

$$r = r^* - \left(\frac{m}{e^2} \right)$$

حيث الرمز r = معامل الثبات بعد تصحيح أثر السرعة

r^* = معامل الثبات بعد تصحيح سبيرمان - براون

m = متوسط الأسئلة المتروكة

e^2 = تباين الخطأ (تباين الاجابات الخاطئة والاجابات
المحذوفة)

٢ - مدى الفروق الفردية في العينة :

تمثل طبيعة الفروق الفردية في عينة الأفراد التي تطبق عليها الاختبارات
بفرض تقدير معاملات الثبات عاملاً هاما مؤثرا في هذه المعاملات . ويجب
ان نؤكد ان معامل الارتباط عموما يتأثر بمدى الفروق الفردية في الأفراد ،
فاذا انعدمت هذه الفروق بالنسبة لخاصية من الخصائص فان معامل ارتباط
هذه الخاصية بغيرها يصبح صفرا ، ويكون من المستحيل علينا ان نتنبأ بمركز
الفرد في أى خاصية أخرى من معرفة درجته في الاختبار الذى وضعناه .
وانا اخذنا مثالا على ذلك اقل تطرفا وليكن معامل الارتباط بين درجات اختبار
للفهم اللغوى واختبار للاستدلال الحسابى طبقا على عينة متجانسة تجانسا
شديدا (مثلا طلاب احدى كليات الجامعة) فان معامل الارتباط في هذه
الحالة قد يكون منخفضا انخفاضا واضحا دون ان يدل ذلك على ضعف العلاقة بين
الاختبارين . اما اذا اعطينا الاختبارين لعينة غير متجانسة فان معامل الارتباط
بينهما قد يكون أعلى منه في الحالة الأولى .

ويوضح لنا ذلك أثر تحديد مدى الفروق الفردية على معامل الارتباط بوجه عام . ويطلق على هذه الظاهرة في علم الاحصاء اسم المدى المحدود للعينة restricted range ، ويعتمد ذلك في جوهره على مدى التباين في العينة .

وهذا القول يصدق في عمومه على معاملات الثبات - باعتبارها في الأصل معاملات ارتباط - وتوجد معادلات كثيرة لتقدير معامل الثبات المتوقع عندما يزيد أو يقل الانحراف المعياري في الجماعة .

٣ - خصائص العينة :

يختلف معامل الثبات بين الجماعات التي يتفاوت متوسط مستويات القدرة فيها . وهذه الفروق لا يمكن التنبؤ بها أو تقديرها بأي معادلة احصائية، ولكن يمكن اكتشافها وتحديدتها بتجريب الاختبار على جماعات تختلف في العمر أو في مستوى القدرة . ومثل هذه الفروق قد تنشأ من أن المستويات المختلفة لصعوبة الاختبار تقيس في الواقع قدرات مختلفة نوعاً ما بعضها عن بعض ، أو قد تنشأ من الخصائص الاحصائية لنوع الدرجة الذي نستخدمه في تصحيح الاختبار كما هو الحال في درجة نسبة الذكاء في اختبار ستانفورد - بينه مثلاً ، حيث نجد أن معاملات الثبات قد تكون منخفضة نسبياً في الجماعات الأصغر عمراً أو الأقل قدرة لأن هذه الدرجات تتأثر في هذه الجماعات بالتخمين ، ولذلك لا يجب تطبيق الاختبار في هذه المستويات .

ومن الواضح أن كل معامل ثبات يجب أن يصحبه وصف دقيق لنوع المجموعة التي حسب من درجات أفرادها . ويجب أن ننسب على وجه الخصوص إلى تباين القدرة ومستوياتها في العينة . ولا يكون معامل الثبات الذي يتضمنه دليل الاختبار صالحاً إلا للجماعات المشابهة لعينة التقنين . ولذلك يحسن أن نقسم عينة التقنين هذه إلى مجموعات فرعية أكثر تجانساً بالنسبة لعوامل الجنس والعمر والصف الدراسي (أو المرحلة التعليمية) والمهنة وغيرها ، ثم حساب معاملات ثبات منفصلة لكل من هذه المجموعات . وفي هذه الحالة يمكن الاستفادة من معاملات الثبات التي حسبت من عينات التقنين حين نستخدم الاختبار في الأغراض التطبيقية والعملية .

٤ - عدد أسئلة الاختبار :

يزداد معامل ثبات الاختبار مع زيادة عدد الأسئلة التي تكونه ، ولعل هذا العامل لا يحتاج الى مزيد من التفصيل بعد مناقشتنا المستفيضة السابقة لتصحيح اثر تجزئة الاختبار في طريقة التجزئة النصفية ، حيث يفترض أن معامل ثبات نصف الاختبار أقل من معامل ثبات الاختبار الكلى .

٥ - تباين الخطأ :

يؤثر في ثبات الاختبار أى مصدر لتباين الخطأ يكون فعالا في الموقف الاختباري ، ومفهوم الثبات يعتمد في جوهره كما بينا أنفا ، على مدى تحرر الاختبار من هذه المصادر ، وعلى الباحث النفسى أن يجرب كل سبيل ممكن للتأكد من ذلك ولعل طرق حساب الثبات التي عرضناها تهيب له فرصا متعددة متنوعة للوصول الى هذا الهدف .

الخطأ المعياري للقياس :

يمكن التعبير عن مفهوم ثبات الاختبار في ضوء الخطأ المعياري للقياس (ع ق) ، أو الخطأ المعياري للدرجة في الاختبار . ويفيد هذا المفهوم فائدة كبرى في تفسير درجات المفحوصين كأفراد ، ولذلك تفوق أهميته في بعض الحالات مفهوم الثبات نفسه . ويمكن حساب الخطأ المعياري للقياس من معامل الثبات مباشرة بالمعادلة الآتية :

$$\sigma_{\text{ع ق}} = \sigma_{\text{ع د}} \sqrt{1 - r}$$

حيث ع ق = الخطأ المعياري للقياس

ع د = الانحراف المعياري لدرجات الاختبار

ر = معامل الثبات

بشرط أن يحسب ع د ، لنفس العينة .

ولتوضيح ذلك نعطي مثالا : نفرض أن الانحراف المعياري لأحد

اختبارات القدرة الموسيقية ١٠ ومعامل ثباته ٨٤ مر فائقا نحصل على الخطأ المعياري للقياس مباشرة بالمعادلة السابقة كما يلي :

$$عق = ١٠ = \sqrt{١٧ - ٨٤} = ٤$$

ويمثل هذا المقدار الدرجة التي تتذبذب فيها درجات المفحوص اذا افترضنا أنه تعرض للقياس بنفس الاختبار عددا كبيرا من المرات وأن درجاته تتوزع توزيعا اعتداليا ، أى أنه يعبر عن الانحراف المعياري لدرجات المفحوص الواحد في نفس الاختبار ، لو أعيد تطبيق هذا الاختبار عليه عشرات المرات .

لنفرض أن أحد المفحوصين حصل على الدرجة ٤٥ في هذا الاختبار ونريد أن نحدد الحدود الدنيا والعليا التي تتوقع لدرجاته الا تتجاوزها لو أعيد عليه القياس (حدود الثبات) . في هذه الحالة يقوم الباحث بما يلي :

١ - اختيار احدى نقاط المنحنى الاعتدالى ولتكن النقطة التي تزيد على المتوسط (م) بما يعادل + ٢ر٥٨ ع ، وهي النقطة التي تحددها نظرية المنحنى الاعتدالى بأن عندها يمكن أن يتم استيعاب ٩٩٪ من درجات هذا المفحوص لو أعيد تطبيق الاختبار عليه عشرات المرات ، أما النسبة المتبقية الباقية (١٪) فتمثل احتمال حدوث خطأ القياس وعدم حصول المفحوص على الدرجة ، أى أن حدوث احتمال حدوث الخطأ في هذه الحالة $\pm ١٪$.

٢ - حساب النقطة التي تقابل (٢ر٥٨ ع) في درجات المفحوص بافتراض أن الخطأ المعياري (ع ق) هو في جوهره انحراف معياري (ع) وفي مثالنا الحالي تكون هذه النقطة $\pm ٢ر٥٨ \times ٤ = \pm ١١ر٣٢$ والتي قد تكون موجبة أو سالبة حسب موضع الانحراف المعياري من المتوسط .

ومن هذا المثال يمكن أن نستنتج أن درجة هذا المفحوص عند تطبيق الاختبار عليه مرة واحدة سوف تمتد بين (١١ - ٤٥) و (٤٥ + ١١) أى بين ٣٤ ، ٥٦ بدرجة من اليقين يصل الى ٩٩٪ . فاذا طبق عليه هذا الاختبار (التقويم النفسى)

أو أى اختبارات مكافئة مائة مرة فإن درجته التى يمكن أن تقع خارج هذا النطاق لن تحدث الا مرة واحدة ، ويمثل هذا النطاق حدود الدرجة الحقيقية true score للفرد .

ومن الواضح أن الخطأ المعيارى للمقياس يصف الأداء الفردى ويفسره بينما يعبر معامل الثبات عن الأداء الجماعى ، ولهذا لا تصلح الأخطاء المعيارية للمقارنة بين الاختبارات فى درجة الدقة ، وفى هذه الحالة لا تصلح الا معاملات الثبات .

ولعل من أهم فوائد الخطأ المعيارى للمقياس - الى جانب تفسيره لأداء الفرد - أنه يجعل الفاحصين لا يبالغون فى تقدير الفروق الضئيلة بين الدرجات . وهذا الحذر يفيد الفاحص سواء عند المقارنة بين درجات مفحوصين مختلفين فى الاختبار الواحد أو المقارنة بين درجات نفس المفحوص فى اختبارات مختلفة . وكذلك يجب فهم الدرجات التى يحصل عليها المفحوصون بعد المعالجات التجريبية فى ضوء الخطأ المعيارى للمقياس . الا أننا فى هذه الحالة نحتاج الى حساب الخطأ المعيارى للفروق بين المقاييس باستخدام المعادلة الآتية :

$$ع_2 = \sqrt{ع_1^2 + ع_2^2}$$

حيث

ع ف = الخطأ المعيارى للفروق بين مقياسين

ع ق ١ ، ع ق ٢ = الخطأ المعيارى للمقياس الأول والمقياس الثانى على التوالي .

ويمكن إعادة صياغة المعادلة السابقة باستخدام معاملات الثبات كما يلى .

$$ع_2 = ع_1 \sqrt{ر_1 - ر_2}$$

مع ملاحظة أن الرمز ع يستخدم للاختبارين معا لأن درجاتهما يجب أن

تحول الى نفس المقياس قبل المقارنة بينهما (اى يتم تحويل الدرجات الخام الى درجات معيارية او درجات معيارية معدلة على النحو الذى سنبينه فى الفصل السابع) . ويوضح ذلك المثال الآتى :

نفرض أننا طبقنا اختبارين معامل ثباتهما ٩٧ ر ، ٩٢ ر على التسوالى ، وأن درجاتهما الأصلية الخام حولت الى نسب ذكاء انحرافية (درجات معيارية معدلة) متوسطها ١٠٠ وانحرافها المعيارى ١٥ ، يمكن أن نحسب الخطأ المعيارى للفرق بين درجتى هذين الاختيارين بالتعويض فى المعادلة السابقة كما يلى :

$$ع ب = ١٥ \sqrt{\frac{١}{٢} (٨٧ - ٨٣)^2} = ٨ ر١$$

ويمكن أن نستخدم هذه القيمة بنفس المعنى الذى استخدم فيه الخطأ المعيارى للمقياس الواحد .

الفصل الخامس

صدق الاختبارات

معنى صدق الاختبار :

يتعلق موضوع صدق الاختبارات Validity بما يقيسه الاختبار والى أى حد ينجح فى قياسه ويتصل هذا بمدى وصولنا الى تنبؤ دقيق أو استنتاج صحيح من الدرجة التى يحصل عليها المفحوص فى الاختبار . وهذا الموضوع لا يقتصر على عملية القياس وحدها وإنما قد يمتد الى المنهج التجريبي بصفة عامة والذي يتطلب اختبار فروض علمية حول طبيعة الاختبار وما يقيسه .

وإذا كان انصرق مسألة على درجة كبيرة من الأهمية فى تقرير صلاحية أو فائدة أى فرض علمي أو فى تحديد معنى مفهوم معين فانه فى الواقع هو الذى يحدد قيمة الاختبار ومغزاه . ويجب أن نقرر منذ البداية أن عنوان الاختبار لا يدل بالضرورة على ما يقيسه الاختبار لأنه عادة ما يكون مجرد كلمة أو عبارة قصيرة تهدف الى التعرف على الاختبار . ولا يمكن أن نستدل على العملية النفسية التى يقيسها الاختبار الا بفحص مجموعة من المعلومات والبيانات التى تتوافر لنا بالطرق الصحيحة وتكون مستقلة نسبيا عن الاختبار .

طرق تحديد الصدق :

يمكن القول بوجه عام أن جميع طرق تحديد صدق الاختبارات تبحث عن العلاقة التى تربط أداء الأفراد فى الاختبار وبعض الحقائق والمعلومات التى تتصل بالعملية النفسية موضوع الاختبار ، والتى (أى هذه المعلومات والحقائق) يمكن ملاحظتها ملاحظة مستقلة . وهناك كثير من الطرق التى تستخدم فى هذا الصدد ولها مسميات كثيرة ، إلا أننا سوف نتحدث عن الأنواع الأربعة من الصدق التى وصفها كتيب التوصيات الفنية للجمعية الأمريكية لعلم النفس (١٤١٣) والتى عالجتها مؤلفات أنستازى (١١) وكرونباك (٢٠) بالتفصيل .

١ - صدق المحتوى :

يقصد بصدق المحتوى Content Validity أن يقوم الباحث النفسى بفحص مضمون الاختبار فحصا دقيقا منتظما لتحديد ما اذا كان يشتمل على عينة ممثلة لميدان السلوك الذى يقيسه كما هو الحال فى الاختبارات المرجعة الى المعيار أو تحتوى على معظم ان لم يكن جميع جوانب هذا السلوك ، كما هو الحال فى الاختبارات المرجعة الى المحك . ومثل هذا النوع من الصدق نستخدمه فى العادة فى تقويم الاختبارات التحصيلية التى نعدّها لقياس مقدار ما تمكن التلاميذ من اتقانه من جوانب مهارة معينة أو ما حصلوه من مقرر دراسى أو برنامج مهنى . وقد يبدو أن مجرد فحص مضمون الاختبار يكفى للتدليل على صدقه وصلاحيته للفرض الذى وضع من أجله ، فاختبار معين فى الضرب أو التاريخ قد يبدو صادقا بالتعريف اذا اشتمل على أسئلة تدور حول عملية الضرب أو التاريخ ، الا أن ما نحب أن ننبه اليه أن هذا قد لا يكون صحيحا ، فليس مجرد تأمل محتوى الاختبار يدل على صدقه .

ولكن الأمر ليس على هذا القدر من السهولة لأن من أهم المشكلات التى نواجهها هنا هى مشكلة عينة المحتوى أو أصله الإحصائى العام . ف نطاق المحتوى الذى نختبره يجب أن يتم تحليله تحليلا دقيقا منتظما حتى نتأكد من أن جميع الجوانب الأساسية تغطيتها أسئلة الاختبار تغطية ملائمة وبالنسب الصحيحة حتى لا يكون لدينا عدد كبير من الأسئلة فى جوانب معين على حساب الجوانب الأخرى . ويتطلب ذلك بادئ ذى بدء ، وصفا تفصيليا دقيقا لمدى المهام والمثيرات والمواقف والمواد التى يستوعبها نطاق المحتوى ، كما يتطلب تفصيل أنواع الاستجابات التى يتوقع لها أن تصدر عن المفحوص ويجب أن يهتم بها الفاحص أو الملاحظ أو المصحح ، والتى قد تسدل على عمليات نفسية معينة ، وكذلك توصيف الأهداف التعليمية أو التدريبية أو العلاجية بشرط أن تصاغ فى عبارات إجرائية سلوكية تقبل الملاحظة والقياس ، وبهذا يكون صدق المحتوى أكثر اعتمادا على مدى الارتباط بين استجابات المفحوص على أسئلة الاختبار وميدان السلوك الذى نختبره ، وأى تغيير فى أحد هذه العناصر الثلاثة يغير مباشرة فى طبيعة الاختبار . وفى قياس القدرة على القراءة مثلا قد يلجأ الباحث - بالنسبة لأهداف معينة - الى بناء اختباراه فى حدود المادة القرائية التى تقدمها المدرسة . وقد يلجأ - بالنسبة لأهداف

أخرى - إلى أن يتناول الاختبار نطاقا أوسع من المواد المقررة مثل الروايات والصحف والمجلات . ولهذا لا بد لمؤلف الاختبار من أن يحدد في دليل اختباريه وبدقة عناصر التحليل الثلاثة التي أشرنا إليها (المواد والاستجابات والأهداف) حتى يمكن لأي فاحص أن يحكم على مدى ملاءمة الاختبار لأهدافه في التقويم النفسى .

وصدق المحتوى يفيد كما سبق أن قلنا في اختبارات التحصيل ، فهذه الاختبارات يكون لمحتواها - فى العادة - أصل احصائى يمكن أن يعبر عنه كله فى الاختبار ، أو تنتقى منه عينة ممثلة ، وخاصة فى الاختبارات التحصيلية التى تستخدم فى تقويم مدى الوصول الى الأهداف التعليمية من التويات المنخفضة . إلا أن صدق المحتوى لا يكون كافيا لاختبارات الاستعدادات وقاييس الشخصية بل ولاختبارات التحصيل التى تقيس العمليات المعرفية العليا ، بل قد يكون مضللا فى كثير من الحالات . فعلى الرغم من أن هذا النوع الأخير من الاختبارات لابد أن يتوفر فيها محكا الملاءمة والفاعلية بالنسبة لمحتواها وخاصة فى المراحل الأولى لاعدادها إلا أن هذا وحده لا يكفى ، بل لابد من أن نلجأ بعد ذلك الى وسائل التحقق التجريبي المضبوط ، التى سوف نتحدث عنها فيما بعد ، لتقدير صدقها . ويتضح ذلك اذا علمنا أن اختبارات الاستعدادات والعمليات المعرفية العليا والشخصية ، على عكس الاختبارات التحصيلية ، لا تتشابه تشابها داخليا كبيرا مع ميدان السلوك الذى تعد عينة منه ، وبذلك لا يكون لمحتواها أى دور فى التحقق من الفروض التى يصوغها مؤلف الاختبار عند انتقاء محتواه ، لأن هذه الفروض تتطلب فى هذه الحالة تحقيقا تجريبيها منظما حتى يمكن التأكد من صدق الاختبار . وبعبارة أخرى نستخدم الاجراءات التجريبية المنظمة للتحقق من درجة التشابه بين عينة السلوك الذى يتضمنها اختبار الاستعداد والسلوك الذى تنتمى اليه .

ومن ناحية أخرى فان اختبارات الاستعداد والشخصية على عكس الاختبارات التحصيلية لا تعتمد على مقرر دراسى أو على برنامج تدريبي أو على مجموعة موحدة من الخبرات السابقة يمكن منها صياغة محتوى محدد للاختبار ، وانما يتفاوت الأفراد كثيرا فى طرق عملهم وفى العمليات النفسية التى يستخدمونها عند الاجابة على نفس الأسئلة . وعلى ذلك فان الاختبار

الواحد قد يقيس عمليات سلوكية مختلفة في الأفراد المختلفين . وفي مثل هذه الأحوال يستحيل علينا تحديد هذه العمليات التي يقيسها الاختبار من مجرد فحص محتواه . فقد يحل طالب الجامعة مشكلة معينة مستخدما التعبير اللفظي (أو الرمزي عموما ومنه الرياضي) ، بينما يحل الميكانيكي نفس المشكلة ويصل الى نفس الحل مستخدما التصور البصري المكاني . كما أن اختبارا معيننا للاستدلال الحسابي لطلاب المرحلة الاعدادية قد يقيس مجرّه الفروق الفردية في سرعة العمليات العددية عند طلاب الجامعة .

ويجب ألا نخلط بين صدق المحتوى والصدق الظاهري face validity فالصدق الظاهري ليس صدقا بالمعنى العلمي للكلمة لأنه يدل على ما يبدو أن الاختبار يقيسه من الظاهر لا على ما يقيسه بالفعل . كما أنه يعني ما اذا كان الاختبار يبدو صادقا في نظر كل من المفحوصين والفاحصين . والواقع أن موضوع الصدق الظاهري يدخل في باب العلاقة الطيبة التي يكونها الفاحص مع المفحوص ، وهو من الأمور المرغوب فيها دون شك رغم ما يسببه من خلط وتداخل مع مفاهيم الصدق الحقيقية فمما لا شك فيه أنه اذا بدا الاختبار للمفحوصين أو المفحوصين لا معنى له ، أو لا علاقة له بالظاهرة موضوع القياس ، أو غير مناسب أو ساذجا أو سخيفا ، فمن المحتّم أن تعاون هؤلاء سيكون ضعيفا حتى ولو كان الصدق الحقيقي للاختبار عاليا . بعبارة أخرى لابد أن يتوافر في الاختبار هنا الصدق الظاهري حتى يكون أكثر فاعلية في مواقف القياس العملية .

ويمكن أن يتحسن الصدق الظاهري للاختبار بمجرد اعادة صياغة أسئلته بحيث تبدو أكثر ارتباطا بموقف القياس الفعلي . فمثلا اذا كان لدينا اختبار للاستدلال الحسابي البسيط نريد أن نطبقه على مجموعة من المتقدمين للمؤن الميكانيكية يحسن أن نستبدل أسئلة الحساب التقليدية التي تتضمنها الكتب المدرسية بأسئلة أخرى أكثر ارتباطا بالعمل الميكانيكي دون أن يحدث أي تغيير يذكر في الوظيفة السلوكية التي نقيسها . ومع ذلك فإن الصدق الظاهري ذاته لا يمكن أن يكون بديلا لأنواع الصدق الأخرى التي تتحدد بالوسائل الموضوعية ، كما أن تحسين الصدق الظاهري لا يؤدي بالضرورة إلى تحسين الصدق الموضوعي .

طرق تحديد صدق المحتوى :

توجد عدة طرق لتحديد صدق المحتوى نذكر منها ما يلي :

١ - **جدول المواصفات :** يساير هذا النوع من الصدق عملية بناء الاختبارات خطوة خطوة ، لأن اعداد أسئلة أو مفردات هذه الاختبارات انما يتم عن طريق الفحص الدقيق الذي يعتمد على تحليل المحتوى للمثيرات والاستجابات والأهداف (التي فصلناها آنفا) ، واستشارة الخبراء واللجوء الى الحكمين . وعلى أساس المعلومات التي يتم جمعها يتم تحديد مواصفات الاختبار Test specifications والتي تتضمن الموضوعات التي يجب أن يشملها الاختبار ، والاستجابات التي يجب اختبارها . والطريقة الأكثر يسرا في تحديد مثل هذه المواصفات هي اعداد جدول ثنائى يسمى جدول المواصفات table of specifications أو مخطط blue print . ولأن هذا الجدول أكثر شيوعا في الاختبارات التحصيلية نعطي مثالا عليه من هذا الميدان الا أنه لا يوجد ما يمنع من استخدامه في بناء الاختبارات من الأنواع الأخرى .

جدول مواصفات لعينة المحتوى والسلوك (الأهداف) : يلعب جدول المواصفات دورا هاما في الاختبارات المرجعة الى المعيار والاختبارات المرجعة الى الحدك ، والفرق بين هذا الجدول في الحالتين أنه في النوع الثانى من الاختبارات يشتمل على كل الأصل الاحصائى للعناصر الهامة للمحتوى والسلوك (الأهداف) ، اما في النوع الأول فان هذا الجدول يمثل عينة من كل منهما .

ويتم اعداد هذا الجدول في الحالتين عن طريق الفحص الدقيق للمقررات والكتب الدراسية من ناحية ، ونواتج التعلم التي يجب اختبارها (فى ضوء الأهداف) من ناحية أخرى . ثم تحديد الأهمية النسبية للموضوعات والأهداف ، وعلى هذا الأساس يتم تحديد عدد المفردات التي ترتبط بكل هدف فى كل موضوع ، وذلك باعداد جدول ثنائى (جدول مواصفات) توضع فيه الأهداف التعليمية أفقيا والموضوعات رأسيا ، ويعطى الجدول رقم (٣) مثالا على ذلك . ومنه يتضح عدد المفردات التي يشملها الاختبار بالنسبة لكل موضوع من موضوعات المحتوى فى علاقته بكل هدف .

من الأهداف التعليمية (التي قد تكون أهدافا معرفية أو وجدانية أو مهارية حركية) وقد ملئت خانات الجدول على أساس الوزن النسبي لكل هدف (في صورة النسب المئوية الموجودة في السطر الأخير من الجدول) ، والوزن النسبي لكل موضوع (في صورة النسب المئوية الموجودة في العمود الأخير من الجدول) ، وهذه الأوزان النسبية يجب تحديدها مقدما ، والطريقة المتاحة حتى الآن لتحديد الوزن النسبي للأهداف هي استطلاع آراء الخبراء (*) ، أما الوزن النسبي للمحتوى فيمكن تحديده موضوعيا في ضوء الوقت المخصص لتدريسه فعليا ، ويحسن أن تكون في جميع الحالات في صورة نسب مئوية .

وقد ملئت خانات الجدول رقم (٣) في صورة تكرارات متوقعة ، ويقترح هؤلاء أبو حطب (٩) المعادلة الآتية لتحقيق هذا الغرض .

$$Z = \frac{A \times B}{N}$$

حيث يدل الرمز (ز) على الوزن النسبي للخانة .
(أ) على الوزن النسبي للهدف (العمود) الذي تقع فيه الخانة في صورة نسبة مئوية .

(ب) على الوزن النسبي لعنصر المحتوى (السطر) الذي تقع فيه الخانة في صورة نسبة مئوية .

(ن) على المجموع الكلي للأوزان (وهو هنا ١٠٠) .

(*) نحب أن ننبه الى ابتذال طريقة استطلاع آراء الخبراء واللجوء الى المحكّين في عدد كبير من الاختبارات التي ظهرت في السنوات الأخيرة في مصر وبعض البلاد العربية ، الى الحد الذي يدفعنا الى الشك في كثير من بيانات الصديق التي تجمعت منها ، والملاحظ أن كثيرا من الباحثين يلجأون اليها للحصول على معلومات شبيهة امبريقية من أي نوع وبأي سبيل ربما لسهولة الظاهرة ويسرها النسبي ، وفي رأينا أن الباحثين في حاجة الى بذل جهد أكبر في مجال صديق المحتوى لاختباراتهم .

جدول رقم (٢)
جدول الأوزان النسبية لاختبار تحميل موضوع الوحدة في الجغرافيا
بمرحلة التعليم الأساسي

أوزان الأهداف في صورة نسب مئوية			
معرفة	معرفة		مجموع أوزان المحتوي
	مستلزمات	وروز	
أوزان المحتوي في صورة نسب مئوية	٥	٥	٢٢٪
	٦	٨	
	٢	٤	
	٤	٥	
	٢	٢	
٢٠٪		٢٤٪	
مهارات في تفسير الخرائط الجوية	٥	٧	٢٢٪
	٨		
	٣		
	٥		
	٢		
٢٤٪		٢٢٪	
مجموع أوزان المحتوي	٢٢٪	مجموع أوزان المحتوي	٢٢٪
	٢٢٪		
	١٦٪		
	٢٠٪		
	١٠٪		
١٠٠٪		١٠٠٪	

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢٢٪
٦	
٢	
٤	
٢	
٢٠٪	٢٤٪

أوزان الأهداف في صورة نسب مئوية	
معرفة	مجموع أوزان الأهداف
٥	٢

وقد حصلنا على القيم الموجودة في خانات الجدول رقم (٢) من تطبيق هذه المعادلة ، فمثلا في الخانة الأولى من اليمين (معرفة رموز ومصطلحات ٢٢ x ٢٤

الضغط الجوى) حصلنا على الوزن النسبى ٥٪ من ناتج —————
١٠٠

والخانة التى تليها فى نفس السطر (معرفة الحقائق النوعية للضغط الجوى)
٢٢ x ٢٠

من ————— وهى تساوى ٤٠٪ قربناها الى ٥٪ . ويتطلب الامر بالطبع
١٠٠

بعض التقريب حتى نحصل على المجموع الكلى للاوزان وهو عادة ما يكون ١٠٠ ، اذا كانت احكام الخبراء على الأوزان النسبية لاهداف وعناصر المحتوى فى صورة أجزاء من ١٠٠٪ .

وبدل الوزن النسبى لكل خانة من خانات الجدول السابق على النسبة المئوية لعدد المفردات التى يجب أن يشملها الاختبار من النوع الذى تدل عليه الخانة . فمثلا اذا كان اختبارنا يتألف من ٥٠ سؤالاً فيجب أن يكون ٥٪ من هذه الأسئلة من النوع الذى يقيس معرفة رموز ومصطلحات الضغط الجوى ، ١٪ من هذه الأسئلة تقيس فهم العوامل المؤثرة فى الرطوبة والمطر ، ٨٪ منها يقيس المهارة فى تفسير خرائط الرياح ، وهكذا .

ويوضح الجدول رقم (٤) عدد الأسئلة المتوقع من كل نوع فى كل خانة على أساس أن يكون العدد الكلى لأسئلة الاختبار ٥٠ سؤالاً ، واعتماداً على الأوزان النسبية التى حصلنا عليها فى جدول (٣) . فمثلا نتوقع أن يتضمن الاختبار ٣ أسئلة من النوع الذى يقيس معرفة رموز ومصطلحات الضغط الجوى على أساس أن هذا العدد هو ٥٪ من المجموع الكلى لأسئلة

$$٥٠ \times ٥$$

الاختبار أى = ٢.٥ قربناها الى ٣ ، وهكذا .

ويفيد جدول المواصفات في بناء الاختبار وتحديد صدق محتواه من خلال ما يلي :

(١) تحديد عدد المفردات أو الأسئلة التي يتألف منها الاختبار وقد عرضنا طرق تحديد ذلك فيما سبق .

(ب) تحديد نوع المفردات ، فالاختبار من متعدد ليس هو النمط الوحيد مثلا الذي يصلح لقياس جميع الأهداف وجميع جوانب موضوعات المحتوى ، ويمكن للقارئ الرجوع الى الباب الرابع من هذا الكتاب ، والمخصص للاختبارات التحصيلية ليتعرف على طبيعة كل نوع من انواع المفردات وما يقيس من عمليات معرفية ، مع التنبيه خاصة الى ما ينشأ عن بعض هذه الأنواع (وخاصة أسئلة البديلين) من أساليب استجابة response set

(ج) تحديد قالب المهام task format التي يتألف منها الاختبار . ومن المؤكد أن الاختبارات المتشابهة في القالب (اللفاظ أو الصور أو أعمال عيانية محسوسة) والمختلفة في المحتوى (ميكانيكا أو الكترونيات مثلا) تكون فيما بينها أكثر ارتباطا من تلك التي تكون متشابهة في المحتوى (ميكانيكا مثلا) وتعرض في قوالب مختلفة (اللفاظ أو صور أو أعمال عيانية) ، ويقل الارتباط كثيرا بين الاختبارات اذا اختلفت بالطبع في كل من القالب والمحتوى (٢١ : ١٤٦) . وتبدو أهمية هذه المسألة اذا علمنا أن قياس مهارة حركية مثلا باختبار لفظي يكون أقل صدقا من قياسها باختبار عملي .

٢ - بعض الطرق الامبريقية لتحديد صدق المحتوى : وتوجد عدة طرق امبريقية أخرى للتحقق من صدق المحتوى نذكر منها :

(١) المقارنة بين الدرجات التي يحصل عليها المفحوصون في الاختبار ، أو في صورة مكافئة له ، قبل التدريب أو العلاج وبعده للتحقق من التحسن الذي يطرأ على الاختبار نتيجة للتدريب ، مما يؤكد تعلق relevance ما يقيسه الاختبار ببرنامج التدريب أو التعليم أو العلاج النفسي .

(ب) دراسة أنماط الأخطاء الشائعة في الاختبار للتأكد من صدق مفتاح الاختبار .

(هـ) تحليل طرق العمل التي يستخدمها المفحوصون وذلك بإعطاء الاختبار فرديا مع توجيههم الى « التفكير بصوت عال » أثناء حل مشكلات الاختبار .

(د) دراسة أثر بعض العوامل غير المرتبطة بمحتوى الاختبار كالسرعة أو القراءة في علاقتها بما يقيسه الاختبار ، فالسرعة التي لا ترتبط مثلا باختبار يقيس المهارة في الكتابة على الآلة الكاتبة أو اليسر العددي لا ترتبط باختبار آخر يقيس المحصل اللفظي عند الطفل أو قدرته على الاستدلال وحل المشكلات . والقراءة لا ترتبط بقياس المهارة في التصويب ، كما أن قراءة فقرات مطولة لا ترتبط باختبار يقيس القدرة الكتابية (عند الموظفين الإداريين) وإنما تكفي في هذه الحالة قراءة جمل من المستوى الذي يستخدم في الحياة اليومية .

(هـ) تحديد الخصائص اللفظية والكمية لأسئلة الاختبار أو مفرداته ، وهذه العملية هي جزء من عملية تحليل المفردات item analysis أثناء بناء الاختبار ، وتفيد في نفس الوقت - إذا تمت بنجاح - في تحديد صدق محتوى الاختبار .

ويشمل التحليل الكيفي للأسئلة الحكم عليها في ضوء فعاليتها من حيث صياغتها وصلاحيتها ، وفي هذا الصدد قد يلجأ الباحث الى مجموعة من المحكمين لإبداء الرأي في مختلف الجوانب الكيفية للسؤال . أما التحليل الكمي للمفردات فيتضمن على وجه الخصوص تحديد سهولة (أو شيوخ) المفردات ، وتمييزها وتحديد صدقها وثباتها ، وهذه جميعا تفيد في الحكم على مدى ملاءمة محتوى الاختبار ، أو صدق محتواه . وتتناول هذه الجوانب الكمية هنا بشيء من الإيجاز ، وعلى القارئ المهتم بالتفاصيل الرجوع الى فصول (تحليل المفردات) في كتب المقياس النفسي أو الإحصاء النفسي المتخصصة .

معاملات سهولة (أو شيوع) المفردات : يحسب معامل سهولة المفردة
بصفة عامة بالمعادلة الآتية :

$$\text{س} = \frac{\text{س}}{\text{ص} + \text{خ}}$$

حيث س : معامل سهولة المفردة
ص : عدد الاجابات الصحيحة
خ : عدد الاجابات الخاطئة

ويرى فؤاد أبو حطب (١٩) أن هذه المعادلة يمكن أن تكون أكثر
عمومية حين تستخدم لحساب معاملات شيوع المفردات .

$$\text{ش} = \frac{1}{1 + \text{ب}}$$

حيث ش = معامل شيوع المفردة
١ = عدد الأفراد الذين أجابوا على المفردة في الاتجاه (١)
وليكن نعم في استخبارات الشخصية أو صواب في
اختبارات القدرات أو ذكروا المفردة في اختبارات الانحاح .
ب = عدد الافراد الذين أجابوا على المفردة في الاتجاه (ب)
وهو عكس الاتجاه (١) .

وقد استخدم فؤاد أبو حطب مفهوم « معامل الشيوع » Commeness
Coefficient of لأول مرة في منتصف السبعينات في تقنين اختبار رسم
الرجل ، وفي استفتاء للمشكلات الدراسية عند طلاب المرحلة الثانوية .

دقة مفتاح الاختبار وثبات مفرداته : تصلح لحساب ثبات المفردات أو

الأسئلة الطرق المعتادة لحساب الاختيار ككل ومنها إعادة الاختبار إلا أن من أفضل الطرق هنا ما يسمى طريقة الاحتمال المتوالى وهى من اقتراح جتمان ومعادلتها هى :

$$R = \frac{1 - \frac{1}{N}}{\left(\frac{1}{N} - \frac{1}{L} \right)}$$

حيث R = معامل ثبات السؤال .

N = عدد الاحتمالات الاختبارية للمفردة أو السؤال ، ففي أسئلة البديلين N = ٢ وفى سؤال الاختيار من بين ٤ بدائل تكون ٤ وهكذا .

L = الاحتمال المتوالى ، أى أكبر تكرار نسبى لأى احتمال اختياري من الاحتمالات التى تحتوى عليها المفردة أو السؤال

وهذه الطريقة لا ترتبط بالثبات بمعناه المباشر وهى أقرب الى التحقق من دقة مفتاح الاختبار .

تمييز المفردات : يمكن استخدام الفرق بين النسبتين المئويتين للأفراد الذين اجابوا على المفردة من المستويين الأعلى والأدنى كمؤشر على تمييز

المفردة ويرمز له بالرمز (ف) أو (D).

وتوجد عدة طرق أخرى أكثر تعقيدا تستخدم مفاهيم المنحنى الاعتدالى لا يتسع المقام لتناولها بالتفصيل .

الا ان ما يجب ان ننبه اليه ان عددا كبيرا من الاختبارات النفسية لا يتطلب شرط التمييز هذا ، وتظل هذه الاختبارات مع ذلك صادقة من حيث محتواها ، ولعل أشهر الأمثلة المعاصرة على ذلك الاختبارات المرجعية المحك .

مصدق المفردات أو الأسئلة : لعل أهم بيانات تحليل المفردات التى تتصل (التقويم النفسى)

مباشرة بصدق محتوى الاختبار ككل ما يسمى صدق المفردات ، والتي تشمل
ايضا صدق الاختبارات الفرعية Sub-tests التى يتألف منها الاختبار
الكلى .

وتوجد طريقتان لتحديد صدق المفردات أو الأسئلة أو الاختبارات
الفرعية هما :

١ - حساب معامل الارتباط بين درجات المفردة أو السؤال أو الاختبار
الفرعى ودرجات الاختبار الكلى (والتي تعد فى حالة الحالة المحك الداخلى
للصدق) أو درجات أى محك خارجى ، ويتحدد نوع معامل الارتباط المحسوب
تبعاً لطبيعة درجات الجزء (المفردة أو السؤال أو الاختبار الفرعى) من ناحية
و درجات الكل (المحك الداخلى أو المحك الخارجى) من ناحية أخرى .

فإذا كانت درجات الجزء من نوع الكم المنفصل الثنائى أى من نوع الكل
أو لا شيء (مثل نعم - لا ، صفر - ١ ، الخ) ودرجات المحك الداخلى أو
الخارجى من نفس النوع (مثل نجاح - فشل ، تفوق - تخلف) بحسب
معامل الارتباط الرباعى فى هذه الحالة .

أما إذا كانت درجات أحد مكونى العلاقة (الجزء أو المحك) من نوع الكم
المنفصل أو الثنائى ، ودرجات المكون الآخر من نوع الكم المتصل بحسب فى
هذه الحالة معامل الارتباط الثنائى .

أما إذا كانت درجات كل من الجزء (السؤال أو المفردة أو الاختبار
الفرعى) والمحك من نوع الكم المتصل بحسب معامل الارتباط التتابعى لكارل
بيرسون .

٢ - طريقة المقارنة الطرفية : وذلك بتقسيم درجات المحك الداخلى أو
الخارجى الى مستويين ، يستخدم هنا أسلوب أشبه بطريقة المجموعات
المتضادة تتمثل فى انتقاء مجموعتين متطرفتين من المفحوصين فى ضوء
درجاتهم الكلية فى الاختبار (المحك الداخلى) أو فى ضوء درجات المحك
الخارجى ، ويمكن أن تكون هاتان المجموعتان المتطرفتان أولئك الذين ينتسبون

الى الربيع الأعلى (المستوى المرتفع) والى الربيع الأدنى (المستوى المنخفض) بالنسبة للدرجات الكلية فى المحك ، ويرى فلاناجان ان نسب مجموعتين متطرفتين يمكن اختيارهما فى ضوء المحك الداخلى خاصة (اى الدرجة فى الاختبار الكلى) هما أولئك الذين يقعون ضمن نسبة ٢٧٪ من المفحوصين الحاصلين على الدرجات العليا فيه ، والذين يقعون ضمن نسبة ٢٧٪ الحاصلين على الدرجات الدنيا .

وبعد تحديد المجموعتين المتضادتين فى ضوء المحك الداخلى او الخارجى تتم المقارنة بين اداء المفحوصين فيهما على المفردة او السؤال او الاختبار الفرعى ؛ ويستخدم فى هذا الغرض احدى طريقتين . .

١ - فى حالة استخدام نسبة ٢٧٪ العليا ، ٢٧٪ الدنيا يمكن الاعتماد على معاملات الارتباط التى حسبها فلاناجان فى جداوله الشهيرة والتى تعبر عن العلاقة بين معاملات سهولة السؤال او المفردة للمجموعتين العليا والدنيا من ناحية ودرجات المفحوصين فى كل من المجموعتين فى السؤال (وهى من نوع الكم المتصل او الثنائى) وبعبارة اخرى تعبر عن العلاقة بين الدرجة الكلية فى الاختبار ودرجة كل سؤال على جده من أسئلة الاختبار فى المحك الداخلى او الخارجى .

٢ - فى حالة اختيار المجموعتين المتضادتين فى المحك الداخلى او الخارجى فإى طريقة اخرى غير الطريقة السابقة (بطريقة الربيع الأعلى والأدنى ، أو نسبة ١٠٪ العليا ، ١٠٪ الدنيا مثلاً) يمكن حساب دلالة الفروق بين اداء المجموعتين فى المفردة أو السؤال أو الاختبار الفرعى باستخدام احد مقاييس الدلالة الإحصائية المناسب كـ χ^2 أو اختبار «ت» أو تحليل التباين أو غير ذلك من الطرق البارامترية واللابارامترية) .

ونحب أن ننبه هنا الى خطأ شاع فى السنوات الأخيرة فى بعض رسائل الماجستير والدكتوراه ، بل وفى بعض « البحوث » التى يتقدم بها أصحابها الى اللجان العلمية الدائمة للترقية الى وظائف أساتذة وأساتذة مساعدين ، ويتلخص هذا الخطأ فى أن الباحث يختار مجموعتيه المتضادتين فى ضوء المحك ثم يقارن بينهما مستخدماً درجات المحك ذاتها ، وهذا التحليل قد يفيد فى

تدعيم فكرة « تضاد » المجموعتين فحسب ، ولا يتجاوز الأمر الى أبعد من ذلك ، ولا يفيد على أى نحو - بالطبع - فى تقديم أى معلومات عن صدق الاختبار . فالصدق يتطلب - ونكرر - أن يتم تحليل أداء المجموعتين المتضادتين فى المحك على المفردة أو السؤال أو الاختبار الفرعى .

ويمتد أصحاب الخطأ السابق أحيانا الى ما هو أخطر ، فنجدهم يكونون المجموعتين المتضادتين فى ضوء أداء الأفراد فى الاختبار الفرعى ، وليس فى المحك ، ثم يقارنون بين المجموعتين فى درجاتهما فى الاختبار الفرعى نفسه ، باستخدام مقاييس الدلالة الاحصائية ، وبالطبع يحصلون على فروق دالة يستنتجون منها « صدق » الاختبار الفرعى . ليس هذا هو العبث العلمى بعينه ، هو نوع من تحصيل الحاصل ، الذى لا يضيف جديدا الى صدق الاختبار ، ولا يتجاوز حدود « اللعب الاحصائى » السخيف .

(ب) الصدق المرتبط بالمحكات

يدل الصدق المرتبط بالمحكات criterion-related validity على قدرة الاختبار على التنبؤ بسلوك المفحوص فى مواقف محددة أو تشخيص هذا السلوك . ولهذا لابد من الحكم على الأداء فى الاختبارات فى ضوء أحد المحكات . ويقصد بالمحك مقياس مباشر ومستقل لما يهدف الاختبار الى قياسه والتنبؤ به أو تشخيصه ، أو هو ميزان لتحديد مدى صلاحية الاختبار ، أو هو اختبار للاختبار . وعلى ذلك فبالنسبة لاختبار يقيس الاستعداد الميكانيكى يمكن أن يكون المحك أداء المفحوصين فى عمل ميكانيكى فعلى فى مدرسة ثانوية صناعية أو مصنع . وهنا يجب أن نميز بين غرضين يستخدم فيهما هذا النوع من الصدق : أولهما ، التنبؤ الطويل المدى ، وثانيهما التشخيص ، وذلك فى ضوء العلاقات الزمنية بين المحك والاختبار ، فإذا تصاحبت أو تلازمت زمنيا بيانات المحك ودرجات الاختبار يصبح الصدق فى هذه الحالة من النوع « التلازمى » ، أما اذا وجد فاصل زمنى طويل (قد يصل لعدة سنوات) بين درجات الاختبار ومعلومات المحك فإن الصدق يصبح « تنبؤيا » . وهكذا يمكن أن نميز بين نوعى الصدق هذين فى ضوء أهداف القياس . فالمعلومات التى يوفرها الصدق التنبؤى ترتبط بالاختبارات التى تستخدم فى انتقاء الأفراد

وتصنيفهم وتوجيههم تربويا أو مهنيا أو عسكريا وكذلك فى أغراض التنبؤ الكلينيكى . وفى جميع هذه الأغراض تتشابه إجراءات الصدق ، حيث تعطى الاختبار لعينة من الأفراد تمثل المجتمع الأصلي (موضوع الاهتمام) ، ولكن الدرجات التى يحصل عليها أفراد هذه العينة لا تستخدم فى اتخاذ أية قرارات خاصة بهم . كما يجب ألا تعطى هذه الدرجات الى أى شخص أو سلطة يمكن له أو لها أن تتخذ قرارات تؤثر فيهم ، حتى لا يطرأ على المحك أى تشويه أو خلط أو تلوث *criterion contamination* لأن الدرجة الاختبارية فى هذه الحالة قد تؤثر على الوضع النسبى للمفحوص فى المحك ، وخاصة إذا كان من نوع التقديرات . وبعد ذلك تتم متابعة العينة الأصلية حتى تتوافر لدينا عنها المعلومات التى نستقيها من المحك والتى تؤدى بنا الى تحديد المدى الذى تتفق فيه درجات الاختبار مع النتائج التى نلاحظها كما تتمثل فى درجات المحك ، وهذا الاتفاق يحدد لنا مدى تنبؤية الاختبار . وتستخدم لهذا الغرض أساليب احصائية معقدة تعتمد فى جوهرها على أسلوب تحليل الانحدار . . . regression analysis (راجع ١٩)

وعادة ما يتضمن دليل الاختبار قيما خاصة بالصدق التنبؤى مشتقة من استخدام محكات مختلفة بهدف مساعدة من يستخدم الاختبار فى فهم ما يقيسه وذلك على الرغم من أن مستخدمى الاختبارات قد لا يهتمون اهتماما مباشرا بالتنبؤ بأى من المحكات المستخدمة .

أما الصدق التلازمى فهو أكثر ملاءمة للاختبارات التى تستخدم لأغراض التشخيص (*) لا التنبؤ بنتائج المستقبل . وفى رأينا أن الغرض الأخير هو أكثر الأغراض مشروعية فى إطار المرحلة الراهنة من تطور علم القياس النفسى .

ويلجأ الباحثون النى الصدق التلازمى فى بعض الأحيان كنوع من حلول الوسط لمشكلة الصدق التنبؤى وما تتطلبه من وقت طويل ، وفيها تعطى الاختبارات لمجموعة من الأفراد تتوافر لدى الباحث عنهم معلومات خاصة

(*) يستخدم غزاد أبو حطب (٩) مصطلح التشخيص بمعنى أكثر شمولاً من معناه الشائع فى إطار علم النفس الكلينيكى ، ويقصد به هنا وصف الوضع الراهن للسلوك وقد يمتد الى الأحداث الماضية .

بمحك معين ، فنقارن مثلا درجات طلاب الجامعة في اختبار معين بمتوسط تراكمي لدرجاتهم الدراسية حتى وقت الاختبار. كما نقارن درجات العمال في الاختبار بمستوى نجاحهم الراهن في العمل .

وتتمثل اهم مشكلات الصدق التلازمي في صعوبة تحديد علاقات السبب والاثـر ، لأنه يدل بالفعل على علاقة اقتران أكثر من دلالة على علاقة «علية» ولذلك فان استخدامه كبديل للصدق التنبؤي يتضمن صعوبات منطقية ظاهرة ، لأن التنبؤ يتضمن في جوهره هذه العلاقة السببية ، وقد يكون في إنقضاء الفاصل الزمني بين الاختبار والمحك ما يبرر « العلية » في الصدق التنبؤي .

إلا أن هذه الفروق بين نوعي الصدق ربما يتجاوزها ، ما أدخل على مفهوم معامل الارتباط من توسيع في معناه وفي وظائفه ، ويدل على ذلك الاستخدامات المتنوعة لأسلوب « تحليل المسار » path analysis (راجع ١٩) ، فالعلية في هذه الحالة قد تستنتج من البيانات الارتباطية بين الاختبارات والمحكات ، سواء كانت هذه المحكات تنتمي الى الماضي أو الحاضر أو المستقبل .

ومن ناحية أخرى فان للصدق التلازمي استخداماته الخاصة في أغراض معينة فهو أكثر ملاءمة كما ذكرنا للاختبارات التشخيصية والتي تلجأ اليها في حل المشكلات الراهنة أكثر من مشكلات المستقبل . فعند اعداد مقياس للشخصية يهتم الصدق التلازمي بسؤال مثل : هل هذا الطفل عصابي ؟ ، أما الصدق التنبؤي فيجيب على سؤال من نوع آخر وهو : هل من المحتمل أن يصير هذا الطفل عصابيا ؟ ومع ذلك فكثيرا ما تثار التساؤلات حول وظيفة الاختبار في حالة الصدق التلازمي الذي « تتصاحب » فيه معلومات الاختبار ومعلومات المحك ، فالواقع أن الاختبار في مثل هذه الحالة عادة ما يكون بديلا أبسط وأسرع وأقل تكلفة وجهدا من معلومات المحك . فاذا كان المحك يتطلب ملاحظة مباشرة لمريض مدة أسبوعين من لحظة دخوله المستشفى ، فان الاختبار الصادق صدقا تلازميا يمكنه أن يميز الأسوياء والعصابيين والحالات المشكوك فيها في وقت قصير ، وبذلك يختصر لنا عدد الأشخاص الذين يتطلبون هذه الملاحظة الشاملة ، بالإضافة الى اقتصاد الوقت والجهد اللذين يبذلان في الملاحظة المباشرة .

أنواع المحكات : يمكن للاختبار أن يحدد صدقه في ضوء أي عدد من المحكات بعدد الاستخدامات الخاصة لهذا الاختبار . وأي طريقة ملائمة للتقويم النفسى فى أى موقف يمكن أن تعد محكا لغرض معين ، بشرط أن تتوافر فيه شروط الاختبار الجيد التى نتناولها فى هذا الكتاب . لقد اشرنا فيما سبق الى أن المحك هو ميزان للاختبار أو اختبار للاختبار ، ولهذا لا بد من التحوط والحذر فى انتقاء المحكات حتى لا توزن الاختبارات بموازين فاسدة أو تختبر باختبارات مضللة . ونعرض فيما يلى لأنواع المحكات الشائعة .

١ - **التحصيل المدرسى الأكاديمى العام :** هو أشهر المحكات وخاصة بالنسبة لاختبارات الذكاء ، ولذلك توصف هذه الاختبارات عادة بأنها مقاييس الاستعداد المدرسى . ومن مؤشرات هذا التحصيل المدرسى مجموع درجات الاختبارات التحصيلية ، ونتائج النقل وتقديرات التخرج ، ومراتب الشرف ، والجوائز أو المنح ، وتقديرات المعلمين لتحصيل تلاميذهم ، ويستخدم هذا المحك - الى جانب اختبارات الذكاء كما اشرنا - فى حساب صدق بطاريات الاستعدادات الفارقة واختبارات القدرات المتخصصة ، وبعض مقاييس الشخصية ، سواء كان غرضنا من استخدام هذه الاختبارات هو التنبؤ أو التشخيص . وهو محك شائع فى جميع المراحل التعليمية وفى جميع أنواع التعليم .

٢ - **الأداء فى برنامج معين :** هذا المحك أكثر ارتباطا بحساب صدق بطاريات الاستعدادات الفارقة واختبارات القدرات المتخصصة ، ومقاييس بعض سمات الشخصية ، وهو يعتمد على مؤشرات أداء المفحوصين فى برنامج معين من التدريب المتخصص . وفى حساب صدق اختبار معين للقدرة الميكانيكية يمكن الاعتماد على التحصيل النهائى فى مقرر معين يتعلق بأحدى المهن الصناعية . وكما يمكن أن نعتبر التحصيل فى مقررات التعليم التجارى، مثل الاختزال والآلة الكاتبة ومسك الدفاتر ، محكات لاختبارات الاستعدادات فى هذه الميادين . وكذلك فإن أداء طلاب المعاهد الموسيقية والفنية قد يعتبر محكا لحساب صدق اختبارات الاستعدادات الموسيقية والفنية . وقد حسب صدق معظم اختبارات الاستعدادات المهنية باستخدام محكات التحصيل فى كليات الحقوق والطب والصيدلة والهندسة وغيرها . أما إذا كانت الاختبارات من النوع الذى نستخدمه فى أغراض برنامج معين للانتقاء ، فقد

تستخدم سجلات التدريب كمصدر للمعلومات التي نحتاج اليها ، كما هو الحال في حساب صدق اختبارات امتقاء الطيارين ، فالمحك هنا هو أداء الطيارين في مراقف التدريب الفعلية .

ومن المؤشرات الخاصة التي يتضمنها محك الأداء في برامج التدريب يمكن أن نذكر الاختبارات التحصيلية المتخصصة التي نطبقها عقب الانتهاء من برنامج التدريب أو تقسيمات المدربين أو النجاح والفشل في التدريب (الاستمرار والاكمال في مقابل الانقطاع والاستبعاد) . وقد نستخدم درجات الامتحانات في مقررات بذاتها في المدرسة الثانوية أو الكليات الجامعية والمعاهد العليا وذلك لتحديد صدق بطاريات الاستعدادات ، فمثلا يمكن مقارنة درجات الطلاب في اختبارات الفهم اللغوي بدرجاتهم في مقررات اللغة العربية ، ودرجات التصور البصري المكاني بدرجات مقرر الهندسة ، الخ .

ولابد أن نميز في هذا الصدد بين ما يسمى المحكات المتوسطة والمحكات القصوى أو النهائية (١٦) بالنسبة لمؤشرات سجلات التدريب . فعند اعداد اختبار لقياس الاستعداد للطيران أو لقياس الاستعداد الطبي أو الاستعداد للتدريس مثلا نجد أن المحكات النهائية هي الأداء الفعلى في ميدان القتال ، أو النجاح في ممارسة الطب ، أو النجاح في الممارسة الفعلية للتدريس داخل حجرة الدراسة ، ويتطلب ذلك دون شك وقتا طويلا حتى يمكننا الحصول على معلومات من هذه المحكات ، بل انه من المشكوك فيه امكان وصولنا الى محك نهائى في الواقع . بالاضافة الى انه ، حتى اذا توفر لنا مثل هذا المحك النهائى ، فانه قد يتعرض لعوامل كثيرة يصعب التحكم فيها مما يجعل المحك في النهاية لا فائدة منه . فمن الصعب مثلا أن تقوم درجة النجاح النسبى للأطباء الذين يمارسون مختلف التخصصات الطبية في أجزاء مختلفة من الوطن سواء داخل المستشفيات أو عياداتهم الخاصة . ولهذا السبب نلجأ في العادة الى المحكات المتوسطة التي ذكرنا أمثلة لها فيما سبق مثل سجلات الأداء في مرحلة معينة من مراحل التدريب .

٣ - الأداء على العمل نفسه : هو أكثر المحكات صلاحية ، لأنه يعتمد على سجلات تتبعية للأداء الواقعى والفعلى للأفراد في العمل . وقد استخدم العلماء هذا المحك الى حد ما في حساب صدق اختبارات الذكاء العام

والشخصية ، والى حد كبير فى حساب صدق اختيارات الاستعدادات الفارقة والقدرات المتخصصة . ومن المؤكد أن « الأعمال » تختلف فى مستوياتها وفى أنواعها . ويصدق هذا على الأعمال الصناعية والتجارية والمهنية والعسكرية وغيرها ، ورغم أن معظم مقاييس أداء « العمل » قد لا تعتبر محكات نهائية إلا أنها أفضل بكثير من محك سجلات التدريب الذى تحدثنا عنه آنفا . ومع ذلك فقد لا يسمح قياس أداء العمل بتوحيد وتقنين الشروط التى تتوافر أثناء التدريب ، كما أنه يتطلب تتبعاً طويلاً المدى مما يؤدى إلى قلة عدد المفحوصين . وبالتالي يؤثر فى النتائج على نحو لا حيلة للمباحث فيه .

والمؤشرات التى نستخدمها فى قياس درجة النجاح فى « العمل » عديدة منها مقدار الانتاج ، ونوعه ، والاستهداف للحوادث ، والمرتبات والعلوات والترقيات والاستمرار فى العمل وعدم الانتقال إلى غيره ، وطول مدة الخدمة ، ومعدل النمو فى المهنة ، وتقديرات المشرفين والرؤساء للكفاءة المهنية للمفحوص وقد يعتمد مقياس المحك على ملاحظة عينة صغيرة محددة من أداء الشخص فى العمل ، أو قد يشتق من سجل تتبعى للانتاج ، أو تاريخ الشخص فى العمل . بحيث يصبح محكاً شاملاً للفترة التى يقضيها المفحوص فى عمله .

٤ - المجموعات المتضادة : محك « المجموعات المتضادة » ، Contrasted groups ' هو من نوع المحكات المركبة التى تدل على الآثار المتراكمة وغير المضبوطة للحياة اليومية ، أى أنه أقل وضوحاً وأكثر تعقيداً من المحكات الأخرى التى سبق مناقشتها ، ويعتمد فى أساسه على بقاء الفرد واستمراره داخل مجموعة معينة لفترة طويلة نسبياً من الزمن أو عدم بقائه فيها وخروجه منها . وفى حسابنا لصدق اختبار الذكاء يمكن أن نقارن درجات ضعاف العقول المقيمين بأحدى المؤسسات بالدرجات التى يحصل عليها تلاميذ مدرسة عادية فى نفس السن تقريباً . وفى مثل هذه الحالة يتأثر المحك بالعوامل العديدة التى تتحكم فى هذه المؤسسات بوجه عام وتجعلها بيئة مختلفة عن المدرسة العادية . وكذلك إذا أردنا حساب صدق اختبار الاستعداد الموسيقى أو الاستعداد الميكانيكى فيمكن أن نقارن بين الدرجات التى يحصل عليها الطلاب الذين يدرسون الموسيقى أو الهندسة فى كليتين أو معهدين متخصصين ، وبين درجات طلاب آخرين يدرسون فى كلية جامعية لا تعد طلابها أعداداً منظماً فى هذين المجالين .

ويمكن للباحث الذى يلجأ الى هذه الطريقة فى حساب الصديق أن يكون المجموعات المتضادة مستخدما ما يشاء من المحكات ، مثل الدرجات المدرسية أو تقديرات النجاح والفشل أو الأداء فى العمل أو غيرها ، وذلك باختصار الحالات المتطرفة فى توزيع درجات المحك . وتصبح المجموعات المتضادة فى هذه الحالة هى جماعات تمايزت تدريجيا بفعل مطالب الحياة اليومية وخبراتها العديدة .

وتستخدم طريقة المجموعات المتضادة فى حساب صديق اختبارات الشخصية على وجه الخصوص ، فمثلا يمكن أن نقارن أداء من يمارسون أعمالا قيادية أو اقناعية بأداء من يمارسون أعمالا كتابية فى اختبار معين للسمات الاجتماعية ليتمكن حساب صديق هذا الاختبار ، مفترضين أن الأفراد الذين يعملون بمهن معينة تتطلب التعامل الاجتماعى (البيع - التدريس . الخ) وظلوا يقومون بها لفترة من الزمن سوف يتقدمون على غيرهم من الأفراد الذين يعملون فى ميادين أخرى لا تتطلب مثل هذا النوع من التعامل كالأعمال الكتابية والهندسية . بل يمكن أن نقارن بين أداء الطلاب الذين يقومون بنشاط كبير خارج المنهج وأداء غيرهم من الطلاب الذين لا يقومون بهذه الأنشطة خلال فترة دراسية معينة . كما أننا كثيرا ما نلجأ الى الجماعات المهنية لحساب صديق اختبارات الميول (ومنها الميول المهنية) ومقاييس الاتجاهات . ومن الجماعات التى تستخدم فى حساب صديق مقاييس الاتجاهات الجماعات السياسية أو الدينية أو غيرها من الجماعات الخاصة التى تمثل وجهات نظر مختلفة اختلافا واضحا حول موضوعات اجتماعية معينة . وقد تستخدم جماعات معينة مثل الأحداث المنحرفين المودعين فى المؤسسات فى مقابل الأحداث غير المنحرفين ، أو العصاة فى مقابل الأسوياء ، عند حساب صديق اختبارات الشخصية التى تتعلق بالتكيف الانفعالى أو الاجتماعى . وفى بعض هذه الاختبارات يلجأ الباحثون الى التشخيص الطبى النفسى باعتباره أساسا لاختيار الأسئلة ودليلا على صديق الاختيار ، إلا أن التشخيص الطبى النفسى لابد أن تتوفر عدة شروط هامة ليصبح محكما مفيدا ، من أهمها أن يستند الى ملاحظة طويلة المدى ودراسة مفصلة لتاريخ الحالة ، أما اذا اعتمد على مجرد مقابلة شخصية قصيرة أو على فحص طبى نفسى غير عميق قلن يكون هناك سبب يدعونا الى توقع أن يكون مثل هذا التشخيص أفضل من الدرجة التى يحصل عليها الفرد فى الاختبار موضع البحث ، بعبارة أخرى

يصبح المحك المفترض (التشخيص الطبى النفسى) كالاختبار الجديد نفسه ، مجرد مؤشر يحتاج هو فى ذاته الى مزيد من البحث حول صدقه ولا يكون محكا كافيا نعتمد عليه فى تقدير صدق الاختبار .

٥ - التقديرات : سبق أن تحدثنا عن التقديرات ratings التى يعطيها المعلمون لتلاميذهم أو المشرفون على العمال لعمالهم وغيرهم عند حديثنا عن المحكات الأخرى والتى نطلبها بهدف الحصول على معلومات تتصل بهذه المحكات مثل التحصيل المدرسى الأكاديمى أو الأداء فى نوع متخصص من الأعمال أو النجاح المهنى . إلا أن التقديرات قد تصبح محكا نعتمد عليه فى ذاته عند حساب الصدق حين تتطلب حكما شخصيا من جانب الملاحظ (المعلم أو المشرف) على الخاصية النفسية التى يقيسها الاختبار . ومن ذلك أن نطلب من الملاحظين (المعلمين أو المشرفين) أن يعطوا تقديرات للمفحوصين فى خصائص معينة تقيسها الاختبارات مثل السيطرة أو القيادة أو الأصالة أو الابتكار أو الأمانة . الخ . هذه التقديرات تعد ذاتها محكات واستخدمها الباحثون فى حسابهم لصدق كثير من الاختبارات ، وبخاصة اختبارات الشخصية، وذلك لصعوبة الحصول على محكات موضوعية فى هذا الميدان وبخاصة بالنسبة للسمات الاجتماعية التى لا يمكن الحكم عليها إلا من خلال مثل هذه التقديرات المعتمدة على العلاقة الشخصية .

والواقع أن التقديرات التى تحصل عليها من ملاحظين لديهم التدريب والخبرة وفى ظروف مضبوطة ضبطا تجريبيا دقيقا توفر لنا معلومات هامة عن المحك . وما يزيد ثقتنا فى هذا المحك أن نحصل على تقديرات مستقلة من أكثر من ملاحظ وحساب معاملات الموضوعية (داجع الفصل الرابع) حتى يمكن التغلب على العوامل الذاتية التى تؤثر فى ثبات هذه التقديرات . كما يمكن أن تزداد دقة التقديرات إذا استخدمنا مقاييس التقدير rating scales بحيث تتضمن وحدات محددة تحديدا واضحا مع وجود ضمانات خاصة للتخلص من أخطاء التقدير الشائعة مثل أثر وقع الهالة halo effect (وهى عبارة عن ميل الملاحظين أو الفاحصين الى التأثر بسمة أو سمات غير مطلوبة ، وتكون فى صالح المفحوص أو فى غير صالحه وتؤدى الى التأثير فى أحكامهم وتقديراتهم للسمة موضع القياس) ، كما لابد أن تتوفر للمقدرين « ألفة بالسمات موضوع القياس » trait acquaintance وفهم لعناهما

بحيث يمكنهم الحكم على وجودها بدرجة معينة لدى المفحوصين . بعبارة أخرى لا يكفي أن يعرف الفاحص المفحوص فترة طويلة ليصبح قادرا على اصدار حكم عليه وإنما يجب أن تكون قد تهيأت له فرصة ملاحظته في مواقف مختلفة تظهر فيها السمة السلوكية موضوع الاهتمام . وفي جميع الحالات لا يجب أن يعطى الفاحصون تقديرات للمفحوص في سمة ليس لديهم عنها معرفة كافية .

وتلعب التقديرات دورا هاما في مقاييس المحك الذي يعتمد على التقييم الكلينيكي . وفي مثل هذه الأحوال لابد من تدعيم الحكم الكلينيكي ببعض المعلومات الاضافية مثل التاريخ المفصل للحالة ودرجات الاختبارات الأخرى وغيرها من الوسائل الموضوعية، أما التقييم النهائي فيعتمد على عملية الحكم كما هو الحال في جميع التقديرات .

٦ - معاملات الارتباط بين الاختبار وغيره من الاختبارات :
من وسائل تحديد صدق الاختبار حساب معاملات الارتباط بين درجات الاختبار التجريبي (الجديد) وبين درجات اختبارات أخرى قائمة بالفعل تقيس ما يفترض الاختبار الجديد قياسه . وقد نلجأ الى هذه الطريقة اذا كان الاختبار الجديد صورة مختصرة أو مبسطة لاختبار موجود فعلا ، وفي هذه الحالة يكون الاختبار الأصلي محكا . كما يمكننا أن نقدر صدق اختبار لفظي (ورقة وقلم) باستخدام اختبار يدوي على مطول يستغرق وقتا طويلا وله معامل صدق مرتفع تم تحديده من قبل . وكذلك يمكن تقدير صدق اختبار جماعي باستخدام اختبار فردي (مثل اختبار ستانفورد - بينيه الذي استخدم كثيرا كمحك لحساب صدق الاختبارات الجماعية للذكاء) . وفي مثل هذه الأحوال يمكن اعتبار الاختبار الجديد في أحسن الحالات صورة تقريبية للاختبار الأصلي . ويجب أن نلاحظ أنه ما لم يكن الاختبار الجديد بديلا مبسطا أو مختصرا للاختبار الأصلي يصبح لا قيمة له ، لأنه لن يكون في أحسن حالاته الا تكرارا لاختبار موجود بالفعل .

بعض مشكلات الصدق المرتبط بالمحكات : لعل من أهم المشكلات التي تعترض نظرية القياس النفسي بالنسبة للصدق المرتبط بالمحكات ما يتصل بمدى عمومية هذا النوع من الصدق ، وخاصة أن أهم ما يميزه أنه نوع من

الصدق العملى للاختبار فى موقف معين بذاته ، كان يحدد فالية الاختبار
لبرنامج تدريبي معين .

وقد بدأت المشكلات فى الظهور حين لوحظ أن معاملات الارتباط بين
اختبارات الاستعدادات المقتنة والأداء فى أعمال يفترض فيها أن تتشابه مع
ما تقيسه هذه الاختبارات ، وخاصة فى بحوث الصدق فى الأعمال الصناعية
لوحظ أن معاملات الصدق اختلفت اختلافات بينة ، كما حصل الباحثون على
ما يشبه ذلك فى بحوث الصدق التى استخدمت التحصيل فى مختلف المقررات
الدراسية محكا ، وأدى ذلك الى سيطرة اتجاه لدى مستخدمي الاختبارات حتى
منتصف السبعينات خلاصته أن صدق الاختبار لا يقبل التعميم الى مواقف
مختلفة . وسادت فكرة « الخصوصية الموقفية » ، واعتبرت فى نفس الوقت
عائقا لاستخدام الاختبارات المقتنة فى انتقاء الافراد . الا أن الدراسات
السيكومترية التى أجريت طوال العقدين الماضيين أكدت أن هذه الاختلافات
فى معاملات الصدق المرتبط بالمحكات هى «محض اصطناع احصائى» ناتج عن
عن صغر حجم العينات المستخدمة فى حساب الصدق ، وعدم ثبات المحك ،
وضيق مدى الفروق الفردية فى عينات التفتين ، وهذه جميعا تؤدي الى اختلاف
معاملات الصدق المستخدمة فى حساب الصدق ، وعدم ثبات المحك ، وضيق مدى
الفروق الفردية فى عينات التفتين ، وهذه جميعا تؤدي الى اختلاف معاملات
الارتباط من ناحية ، والى خفضها من ناحية أخرى .

وقد أمكن التغلب على هذه المشكلة فى السنوات الأخيرة من خلال نموذج
جديد اقترحه شميدت Schmidt وهانتر Hunter واطلقا عليه اسم
«النموذج البايزى لتعميم الصدق» Bayesian Validity Generalization Model
وطبقاه على عينات كثيرة اختيرت من تخصصات مهنية كثيرة وأكد أن صدق
اختبارات الاستعدادات اللفظية والعديدية والاستدلالية يمكن أن يعمم الى عدد
كبير من المهن ، وأن الاختلافات السابقة لم تكن الا محض « مصادفة » .

(د) صدق التكوين الفرضى :

يقصد بصدق التكوين الفرضى Construct validity مدى قياس
الاختبار لتكوين فرضى أو مفهوم نفسى أو سمة . ومن أمثلة هذه التكوينات

الفرضية الذكاء والفهم الميكانيكى والطلاقة اللغوية والعصابية والقلق .
وغيرها . وكل من هذه المفاهيم أو التكوينات يستنتج من أساليب الأداء كما
يفسر حدوث الاتساق الملحوظ فى الاستجابة (٢١) . ويعتمد هذا النوع من
الصدق على وصف أوسع ويتطلب معلومات أكثر حول الخاصية السلوكية
موضوع القياس نحصل عليها من مصادر مختلفة . وإى معلومات تفيد فى
القاء الضوء على المفهوم الذى نحاول قياسه وعلى العوامل التى تؤثر فيه
تضاف الى رصيد صدق الكوين الفرضى .

وتوجد أساليب كثيرة يمكن أن تسهم فى تحديد هذا النوع من الصدق ،
وهى ما يأتى :

١ - تمايز العمر والتغيرات الارتقائية : يمكن أن نلجأ الى هذه الطريقة
فى حسابنا لصدق اختبارات ذكاء الأطفال (كاختبار ستانفورد - بينيه ، أو
معظم اختبارات الذكاء لأطفال ما قبل المدرسة) ، وذلك لتحديد ما اذا كانت
الدرجات تتزايد بتقدم العمر مفترضين أن القدرات تنمو مع العمر خلال
الطفولة وبالتالي نتوقع زيادة فى درجات الاختبار اذا كان الاختبار صادقا .
ويعود الفضل الى بينيه فى صياغة هذا المفهوم الخاص بتحديد المقياس العمرى
للذكاء .

ومن الواضح أن محك تمايز العمر لا يمكن استخدامه فى الوظائف
السلوكية التى لا تظهر تغيرات مطردة مع تقدم العمر ومن ذلك مقاييس
الشخصية بوجه عام . كما أننا يجب أن نشير الى أن هذا المحك محك سلبى
وليس محكا ايجابيا . بعبارة أخرى اذا لم تستطع درجات الاختبار أن تتزايد مع
تزايد العمر فإن ذلك قد يدل على أن الاختبار غير صادق فى قياسه للقدرات
التى يدعى أنه عينة لها ، الا أن تحسين الاختبار بحيث يصبح مقياسا متزايدا
مع تزايد العمر لا يحدد لنا فى ذاته الميدان الذى يقيسه الاختبار . فمثلا يظهر
كل من الطول والوزن زيادة منتظمة مع تزايد العمر ومع ذلك لا يمكن أن نعتبر
أحدهما أو كليهما من مقاييس الذكاء . وكذلك يجب أن نؤكد أنه فى الاختبار
النفسى الذى يقيس خصائص سلوكية تتزايد مع العمر إنما تتأثر هذه الزيادة
أو النمو بظروف البيئة التى يقنن فيها الاختبار ، فالظروف البيئية والحضارية
قد تستثير بعض الخصائص السلوكية وقد تعطل البعض الآخر . ومعنى ذلك

أن محك تمايز العمر لا يمكن أن نعتبره محكا شموليا بحال من الأحوال ، وإنما هو كغيره من المحكات محكوم بالموقف الحضارى الخاص الذى نشته منه .

٢ - معاملات الارتباط بالاختبارات الأخرى : وقد نلجأ فى تقدير صدق التكوين الفرضى الى حساب معاملات الارتباط بين الاختبار التجريبي وبين الاختبارات التى تتشابه معه نظريا فى العمليات التى يقيسها . ويسمى كامبل Campbell وفيسك Fiske الصدق التقاربى converant validity ومن أمثلة ذلك معامل ارتباط اختبار الاستدلال الكمى بدرجات المفحوصين فى اختبار التحصيل الرياضى . إلا أن هذا النوع من معاملات الارتباط لا يكفى ولذلك يقترح كامبل وسك ضرورة أن يضاف الى ذلك حساب معاملات الارتباط بين الاختبار واختبارات أخرى يفترض نظريا أنها تختلف عنه ، ويسمى هذا النوع من الصدق باسم الصدق التمييزى discriminant validity

ويستخدم الصدق التمييزى لمعرفة مدى تحرر الاختبار التجريبي من أثر العوامل التى لا علاقة لها بالعملية السلوكية موضع القياس ، فمثلا اذا حصلنا على معامل ارتباط صفري بين اختبار معين للاستعدادات واختبارات الفهم اللغوى فإن ذلك قد يعنى أن هذا الفهم اللغوى لا يرتبط بالاستعداد موضوع القياس فى الاختبار الجديد وكثيرا ما تتضمن كتيبات الاختبارات معاملات الارتباط المنخفضة بين الاختبار التجريبي واختبارات أخرى تقيس عمليات مختلفة وهى مؤشرات على الرغم من أنها غير مباشرة وسلبية إلا أنها تدل فى نفس الوقت على صدق الاختبار ، لأنه اذا كانت هذه المعاملات مرتفعة كان ذلك مدعاة للشك فى صدق الاختبار إذ تعنى أن الاختبار يقيس عمليات أخرى لم توضع فى الحسبان . إلا أن معاملات الارتباط المنخفضة هذه قد لا تعنى فى ذاتها أن الاختبار التجريبي صادق . ولكي تستخدم طريقة الصدق التقاربى والصدق التمييزى بكفاءة يقترح كامبل وفيسك تصميمًا تجريبيًا خاصًا يسمى « مصفوفة السمات المتعددة والطرق المتعددة »

Multitrait-Multimethod Matrix

ويتطلب هذا التصميم تقويم سمتين نفسييتين أو أكثر بطريقتين مختلفتين أو أكثر ، ثم تحسب جميع معاملات الارتباط الممكنة بين درجات الاختبارات

التي تقيس السمات موضع البحث • ولكي نوضح هذه الطريقة نعطي المثال الافتراضى التالى الذى يقترحه الباحثان •

نفرض أننا أردنا قياس ثلاث سمات نفسية هي (١) السيطرة ، (ب) الاجتماعية ، (ج) دافعية الانجاز ، بثلاث طرق مختلفة هي (١) قائمة من نوع التقرير الذاتى ، (٢) أسلوب اسقاطى ، (٣) بتقديرات الأقران والزملاء ، فإننا نشير فى هذه الحالة الى سمة السيطرة كما تقاس بالتقرير الذاتى بالرمز (١) وكما تقاس بالاختبار الاسقاطى (٢) ، والى سمة دافعية الانجاز كما تقاس بتقديرات الأقران بالرمز (ج) ، وهكذا بالنسبة لجميع البدائل الممكنة •

ويوضح الجدول رقم (٥) مصفوفة ارتباط جميع الدرجات التى يمكن الحصول عليها بهذه الطريقة ، وقد وضعت معاملات الثبات الخاصة بطريقة قياس كل سمة بين قوسين •

وفى هذا الجدول نجد (١) معاملات الصدق فى صورة علاقات بين الدرجات التى حصل عليها المفحوصون لنفس السمة بطرق مختلفة مستقلة (كما هو مفترض فى اجراءات الصدق المعتادة) ، (٢) معاملات الارتباط بين مختلف السمات التى قيست بنفس الطريقة ، و (٣) معاملات الارتباط بين السمات المختلفة التى قيست بطرق مختلفة • ولكى يحصل الباحث على صدق التكوين الفرضى لابد بالطبع ان تكون معاملات الصدق (النوع الأول) اعلى من النوعين الآخرين من معاملات الارتباط • فنحن نتوقع مثلاً أن معامل الارتباط بين دافعية الانجاز كما تقاس بوسائل التقرير الذاتى وكما تقاس بالاختبار الاسقاطى يكون اعلى من معامل الارتباط بين دافعية الانجاز والسيطرة كما تقاس بوسائل التقرير الذاتى ، فإذا حدث أن المعامل الأخير كان مرتفعاً فإننا قد نستنتج من ذلك أن أداء الفرد فى الاختبار أو القائمة المستخدمة يتأثر بعامل مشترك لا يرتبط بما يقيسه (أى نوع من تباين الخطأ الذى تناولناه بالتفصيل فى الفصل الرابع المخصص لموضوع الثبات) ، وقد يكون هذا العامل متصلاً بالفهم اللغوى للأسئلة أو العبارات ، أو رغبة المفحوص فى أن يظهر بصورة موجبة فى جميع السمات (المرغوبة الاجتماعية) •

وتذكر أنستازى (١١ : ١٤٩) أن معامل الثبات فى مصفوفة السمات

جدول رقم (٥) مصفوفة ارتباط متعددة وطرق متعددة
(عن ١١ : ١٤٨)

السمات	الطريقة ١			الطريقة ٢			الطريقة ٣		
	أ	ب	ج	أ	ب	ج	أ	ب	ج
الطريقة ١	(٨٩) ٥١	(٨٩) ٢٧	(٧٦) ٢٨	(٩٢) ٦٨	(٩٤) ٥٨	(٨٤) ٥٩	(٩٤) ٦٧	(٩٢) ٦٧	(٨٥) ٧٨
	١	١	١	١	١	١	١	١	١
	١	١	١	١	١	١	١	١	١
الطريقة ٢	١	١	١	١	١	١	١	١	١
	١	١	١	١	١	١	١	١	١
	١	١	١	١	١	١	١	١	١
الطريقة ٣	١	١	١	١	١	١	١	١	١
	١	١	١	١	١	١	١	١	١
	١	١	١	١	١	١	١	١	١

المتعددة والطرق المتعددة يمثل درجة الاتفاق بين مقياسين لنفس السمة بينهما أقصى حد من التشابه (كما هو الحال فى الصور المتكافئة لنفس الاختبار) ، أما الصدق فيمثل درجة الاتفاق بين مقياسين لنفس السمة بينهما أقصى حد من الاختلاف (كما هو الحال فى درجات الاختبار وتقديرات المعلمين أو الرؤساء) . وطالما أن التشابه والاختلاف بين الطرق هو مسألة نسبية فاننا - من الوجهة النظرية البحتة - يمكن أن نعتبر الثبات والصدق يقعان على مدى متصل واحد . ولعلنا تؤكد فكرة هذا المتصل بين الثبات والصدق اذا ذكرنا أن الثبات فى موجزه يهتم بتباين الخطأ ، بينما الصدق يكاد يتطابق مع مفهوم التباين الحقيقى ، وكلاهما يؤلف التباين الكلى أو الفروق الفردية فى الدرجات التى يحصل عليها المفحوصون فى الاختبار .

ويقترح فسك منذ مطلع السبعينات أن يقوم الباحثون فى مجال الاختبارات النفسية بتحليل مصفوفات عديدة لمعاملات الارتباط بين مقاييس عدد من التكوينات الفرضية (السمات) لطرق مختلفة ، ومن فروضه الأساسية أننا نتوقع فى جميع الحالات أن تكون مقاييس نفس السمة تعطينا نمطا لمعاملات الارتباط يتشابه مع مقاييس سمة أخرى . فمثلا معامل الارتباط بين تقديرات الأقران للسيطرة واستخبار التقرير الذاتى لهذه السمة يظهر على وجه التقريب نفس نمط معاملات الارتباط بين نفس الطريقتين فى قياس سمات المثابرة أو الانتساب أو الاستقلال أو أى سمة أخرى تستخدم فى قياسها هاتان الطريقتان . وقد اختبر هذا الفرض الذى يقترحه فسك فى عدد من البحوث التى أجريت طوال العقدين الماضيين فى مجال مقاييس الشخصية وأظهرت النتائج أن معاملات الارتباط بين السمات فى هذه الحالة لا تظهر الا تشابها جزئيا مع وجود قدر من الفروق الدالة يرجع الى اختلاف الطرق (التباين بين الطرق) بل قد نحصل على أنماط مختلفة لمعاملات الارتباط حتى داخل الطرق (التباين داخل الطرق) . وقد وجدت هذه الفروق مثلا بين وسائل التقرير الذاتى التى بنيت مستقلة واختلفت فى القالب (صواب - خطأ - أو مقارنة ثنائية أو قائمة عبارات) أو فى محتوى المفردات المستخدمة فى الاختبار . ومع ذلك فقد يوجد قدر دال من التباين يمكن قياسه بالطرق المختلفة ويشير الى تكوين فرضى له معنى سيكولوجى ويمكن تعميمه عبر الطرق المختلفة المستخدمة فى تقويمه نفسيا ، وفى هذا وحده معنى صدق التكوين للفرضى .

٣ - التحليل العاملي : قد يكون التحليل العاملي من أهم الوسائل التي تستخدمها في تحديد صدق التكوين الفرضي ، وهو في جوهره أسلوب لتحديد بنية السمات النفسية ، وتقوم على تحليل العلاقات بين البيانات الامبريقية كما تتمثل هذه العلاقات في ضرورة معاملات الارتباط . فنقطة البدء في التحليل العاملي حساب معاملات الارتباط بين درجات كل اختبار والاختبارات الأخرى . التي يجب أن تكون قد طبقت على نفس العينة من المفحوصين . وعندما تفحص الجدول الذي يتضمن هذه المعاملات (والذي يسمى فنيا مصفوفة الارتباط) ونطبق عليه طرقا احصائية خاصة لا يتسع المقام لشرحها قد نلاحظ وجود تجمعات معينة بين الاختبارات مما يوحي بوجود سمات مشتركة بينها . فمثلا اذا كانت الاختبارات التي يرتبط بعضها ببعض ارتباطا عاليا من نوع اختبارات المفردات والمشابهاة والأضداد وتكملة الجمل وكانت معاملات ارتباط هذه الاختبارات بالاختبارات الأخرى منخفضة نستنتج من ذلك وجود عامل احصائي تتشبع به المجموعة الأولى من الاختبارات تشبعا عاليا وتتشبع به المجموعة الثانية تشبعا منخفضا (والتشبع مفهوم احصائي يستخدم في التحليل العاملي ليدل على علاقة المتغير بالعامل) .

وبعد ذلك نبدا عملية تفسير العامل الاحصائي الذي تم الحصول عليه فاذا كانت الاختبارات من نوع الأداء الاقصى كما هو في مثالنا الحالي يفسر العامل بأنه « القدرة على الفهم اللغوي » وقد يفسر العامل بأنه « مهارة » اذا كانت الاختبارات من النوع الذي يسمى « نفسحركي » ، وقد يفسر بأنه « سعة » أو « ميل » أو « اتجاه » أو « قيمة » الخ حسب طبيعة الاختبارات التي تنتمي الى الأداء المميز .

وبالطبع نحن لا نعتمد في هذا التحليل على مجرد تأمل مصفوفة الارتباط والا كانت العملية صعبة ومعقدة وغير مؤكدة بل مستحيلة وخاصة اذا كانت بطارية الاختبارات الكبيرة . ولذلك تستخدم الأساليب الاحصائية الأكثر دقة لتحديد العوامل المشتركة التي يمكن اعتبارها مسئولة عن معاملات الارتباط الناتجة .

وفي التحليل العاملي نلاحظ أن عدد المتغيرات التي نصف في ضوءها أداء كل فرد تختصر باستمرار من عدد كبير من الاختبارات الأصلية الى

عدد أصغر نسبياً من العوامل أو السمات المشتركة . ثم يمكن وصف الفرد في ضوء درجاته في العوامل القليلة العدد . بعبارة أخرى يمكن القول أن من أهم أغراض التحليل العاملي تبسيط وصف السلوك باختزال عدد المتغيرات أي أنه يسير في اتجاه العلم الصحيح من التعدد والكثرة التي تتمثل في متغيرات الاختبارات إلى القلة التي تتمثل في العوامل المشتركة أو السمات .
 إلا أن للتحليل العاملي وظيفة أخرى هي في رأينا أكثر ارتباطاً بمفهوم صدق التكوين الفرضي وهي محاولة اختبار صحة فروض معينة حول البنية العاملية للاختبار الواحد من خلال تحليل مصفوفة ارتباط مفرداته أو أسئلة أو بطارية من الاختبارات النفسية ، وهي وظيفة أشار إليها فؤاد أبو حطب (٩) منذ مطلع السبعينيات ثم ازداد الاهتمام بها في الوقت الحاضر فيما يسمى التحليل العاملي المدعم Confirmatory Factor Analysis .

وبعد تحديد العوامل يمكن أن نقيد منها في وصف التكوين العاملي للاختبار . فكل اختبار يمكن أن يوصف في ضوء العوامل الأساسية التي تحدد درجاته بالإضافة إلى تشعبه بكل عامل . وتشعبات العوامل هذه تمثل بدورها معاملات ارتباط بين الاختبار والعامل . وهذا المعامل يسمى الصدق العاملي الذي هو في الواقع معامل ارتباط بين الاختبار وبيز ما هو شائع أو مشترك أو عام في مجموعة الاختبارات أو غيرها من المؤشرات الدالة على السلوك . ومن المؤكد أن المتغيرات التي نقوم بتحليلها عاملياً تتضمن معلومات مشتقة من الاختبارات أو غيرها من أدوات التقويم النفسي كالتقديرات والاحكام ، وهذه جميعاً يمكن أن تستخدم في محاولة استكشاف الصدق العاملي لاختبار معين وتحديد السمات المشتركة التي يقيسها ، كما يمكن أن يطبق التحليل العاملي على مصفوفات ارتباط المفردات التي يتألف منها الاختبار الواحد .

٥ - الإجراءات التجريبية : من الوسائل الأخرى التي يمكن أن نلجأ إليها لتحديد صدق التكوين الفرضي ما توفره لنا الإجراءات التجريبية من معلومات حول أثر متغيرات معينة على درجات الاختبار . فإذا كان الاختبار التجريبي (الجديد) معاد لقياس القابلية للقلق مثلاً يمكن أن نطبق هذا الاختبار على مجموعة من الأفراد توضع في موقف يثير القلق (مثل الامتحان في ظروف ضاغطة) ثم يحسب معامل الارتباط بين درجات الأفراد في اختبار

جدول رقم (٥) يوضح أنواع المصدق واستعمالات كل منها

أمثلة			
نوع المصدق	السؤال المطروح	الايجراء	الاستخدام الرئيسي
مصدق المحتوى	هل اكتسب المفهوم مقداراً كافياً من المعلومات أو المهارة من البرنامج التدريبي الذي تعرض له؟	تقارن بين أسئلة الاختبار من الوجهة المنطقية بالاحتواء الذي تفرض أن أسئلة الاختبار تنسب إليه .	اختبارات التعميل الدراسي والكفاءة المهنية .
المصدق المرتبط بالمحكات	هل تتبنا درجات الاختبار بأداء معين له أهمية في المستقبل ؟	اعطاء الاختبار ثم نحصل على مقياس لمحك بعد فترة زمنية ثم نقارن الاختبار بالحك	الاختبارات التي تستخدم في الانتقاء والتصنيف .
التنبؤي المصدق	هل تشخص درجات الاختبار صعوبات الأداء في الوقت الحاضر ؟	نمطي الاختبار ثم نحصل على مقياس مباشر دون فاصل زمني ثم نقارن الاختبارات بالحك .	الاختبارات التي تستخدم في أغراض التشخيص التربوي أو الكلينيكي .
التلازمي المصدق	ما المبنى النفسي للمفهوم الذي يقيسه الاختبار ؟	هيمنة الفروض حول الاختبار ثم التحقق منها بالاجراء التجريبي أو الاحصائي المناسب .	الاختبارات التي تستخدم في أغراض البحث العلمي حول الاختبارات النفسية .
مصدق التكوين الفردي			دراسة اختيار القدرة الموسيقية لتعديد إلى أي حد تتأثر الدرجات فيه بموأسل التدريب الموسيقي .
مصدق التنبؤي			مقارنة أداء المتخالفين عقلياً بالأساليب في اختبار الذكاء .
المصدق المرتبط بالمحكات			مقارنة اختبارات القبول بالكليات أو المعاهد بتحصيل الطلاب فيما بعد .
مصدق المحتوى			فحص اختبار القدرة على الاختزال لمعرفة هل يحتوي مضمونه بالفعل على مهارات العمل المكتبي .

القلق قبل التجريب وبين مختلف المؤشرات الفسيولوجية وغيرها التي تدل على القلق خلال التجريب (موقف الامتحان) وبعده . ويمكن القول ان هذا الاجراء التجريبي يشبه الى حد كبير ما تلجأ اليه في تقدير الصدق التنبؤي . باستثناء الدور الايجابي للباحث في تقديم مثير القلق والتحكم فيه . اما المقاييس المباشرة الدالة على القلق فيمكن اعتبارها محكا فراجع عليه تنبؤاتنا من درجات الاختبار .

وقد تفيد الوسائل التجريبية أيضا في تحقيق الفروض المختلفة حول الاختبارات ، ففي مثالنا السابق ذكره يمكن تطبيق اختبار القلق قبل التجريب (اعطاء خبرة القلق) وبعده ثم نتحقق مما اذا كانت درجات الاختبار ترتفع ارتفاعا له دلالة عند اعادة الاختبار وبالطبع يمكن استخدام تصميمات تجريبية تعتمد على المجموعات الضابطة والتجريبية ، فاذا كانت النتيجة ايجابية فان ذلك يدل على ان درجات الاختبار تعكس المستوى الراهن للقلق ، وتبرهن على صدق التكوين الفرضي الذي يقيسه .

العوامل المؤثرة في الصدق :

اذا صيغ الصدق في صورة معامل ارتباط فانه يصبح مثل معامل الثبات يتأثر بالعوامل المختلفة التي تؤثر في معامل الارتباط وخاصة ما يتصل بتجانس العينات والتي تناولناها بالتفصيل في الفصل السابق الا ان هناك عوامل تؤثر في الصدق تأثيرا خاصا ولا بد من ان توضع في الاعتبار .

١ - طبيعة عينة التقنين : من المهم لفهم طبيعة معامل الصدق وصف طبيعة عينة التقنين وصفا مفصلا ، لان الاختبار الواحد قد يقيس عمليات نفسية مختلفة اذا طبق على عينات تختلف في العمر او الجنس او المهنة او المستوى التعليمي او أي متغير آخر يرتبط بمفهوم الصدق . وعلى ذلك فان الاختبار الواحد قد يكون له صدق مرتفع في التنبؤ بمعك معين في بعض العينات ، بينما ينخفض صدقه بالنسبة الى عينات اخرى وربما يكون مقياسا صادقا لعمليات نفسية مختلفة في عينتين مختلفتين ، ومن ذلك مثلا اختبار العمليات الحسابية الذي قد يكون مقياسا صادقا للاستدلال الحسابي عند اطفال الصفوف الدنيا في المرحلة الابتدائية ، بينما يقيس السرعة العددية بصدق عند طلاب المرحلة الثانوية .

٢ - طبيعة العلاقة بين الاختبار والمحك : من المعروف أن حساب معامل

الارتباط بين الاختبار والمحك ، إذا استخدمت فيه طريقة كارل بيرسون أو تقريباتها المختلفة ، فإن ذلك يتضمن افتراضا جوهريا وهو أن العلاقة بين المتغيرين خطية linear على امتداد مدى الفروق الفردية فيهما .

وبالطبع فإن الافتراض قد يكون صحيحا في عدد كبير من الحالات ، إلا أنه في بعض الحالات قد لا تكون العلاقة كذلك ، ومن ذلك مثلا حين يتطلب أحد الأعمال حدا أدنى من فهم القراء بعين العاملين فيه على قراءة بعض التعليمات وفي هذه الحالة نجد أنه حالما توافر هذا الحد الأدنى فإن أى زيادة عليه فى القدرة على فهم القراءة يصبح غير مرتبط بدرجة النجاح فى العمل . وتوصف العلاقة بين الاختبار والمحك فى هذه الحالة بأنها غير خطية non-linear

ومن أمثلة ذلك العلاقة المنحنية curvilinear والمثلثة triangle

أو العلاقة التى تدل على اختلاف التباين heteroscedasticity . ومن أمثلة النوع الأخير من العلاقات حين نجد أن الأداء فى اختبار الذكاء يكون شرطا ضروريا ولكنه ليس كافيا للنجاح المدرسى . وفى هذه الحالة نجد التلاميذ من ذوى الذكاء المنخفض يؤدون أداء سيئا فى المدرسة ، ولكننا حين نفحص أداء التلاميذ من ذوى الذكاء المرتفع سوف نجد أن بعضهم يؤدون أداء جيدا فى المدرسة وبعضهم الآخر قد يؤدى أداء سيئا لأسباب تتصل بنقص الدافعية ، أو فقدان الميل أو غيرها ، وفى هذه الحالة نجد تباينا فى التحصيل بين ذوى الدرجات المرتفعة فى الذكاء أكثر اتساعا مما نجده بين ذوى الدرجات المنخفضة فيه .

٢ - اثر ثبات الاختبار فى صدقه : من بديهيات القياس النفسى أن

الاختبار غير الثابت وغير الدقيق لا يمكن أن يكون صادقا . لأن معنى انخفاض ثبات الاختبار أن تباين الخطأ فيه أكبر من التباين الحقيقى ، وبالطبع فإن تباين الخطأ فى درجات المفحوصين لا يمكن أن يرتبط ارتباطا دالا (أى يتجاوز مستوى الصدقه) مع أى محك . وعلى هذا كلما زاد تباين الخطأ يقل معامل الصدق ، لأن الصدق هو التباين الحقيقى للاختبار أما تحرر الاختبار تماما من تباين الخطأ فإن يصل بالثبات الى حد التمام ويصبح التباين الكلى فى هذه الحالة مساويا للتباين الحقيقى ، وفى هذه الحالة النظرية المستحيلة الحدوث وحدها يمكن للثبات أن يساوى الصدق .

الا ان الاختبار النفسى - كغيره من صور القياس - يكون فى العادة مشوبا بقدر من تباين الخطأ ، ولهذا فان من المتوقع فى هذه الحالة الا يزيد معامل صدقه عن معامل ثباته ، لأنه لا يكون منطقيا على أى نحو أن يرتبط الاختبار مع نفسه أو مع ما يماثله أو يشبهه من صور متكافئة (وهو معنى الثبات) ارتباطا أقل من ارتباطه مع اختبارات أو مقاييس أخرى مختلفة عنه هى ما يسمى المحكات (وهو معنى الصدق) ، ولذلك فكثيرا ما يقال أن الحدود القصوى لمعاملات الصدق المحسوبة هى معاملات الثبات والتي عادة ما لا تتجاوزها النتائج الامبريقية للاختبارات .

الا ان هذه العلاقة بين الثبات والصدق قد أساء عرضها فى عدد من البحوث النفسية التى أجريت فى مصر والعالم العربى فى السنوات الأخيرة ، وخلاصة هذا الخطأ أن بعض الكتب الاحصائية باللغة العربية اشارت الى مفهوم مؤشر الثبات *index of reliability* وهو عبارة عن الجذر التربيعى لمعامل الثبات بمصطلح الصدق الذاتى ، وهى تسمية غير دقيقة فى رأى كاتب هذا الفصل ، لأنها تحمل المؤشر ما لايعنيه ، فهو لا ينتمى الى الصدق بمعناه العلمى الدقيق .

وتلقف باحثونا المستعجلون الفكرة ووجدوا فيها تعويضا عن عجزهم عن البحث فى الطرق الصحيحة لتحديد صدق الاختبارات ، فنجد الواحد منهم يحسب ثبات اختباره ثم يحسب الجذر التربيعى لهذا الثبات ويشير اليه بالصدق ، وكفى الله المؤمنين شر القتال . لون آخر من ألوان العبث العلمى الذى يشهده المسرح المعاصر لعلم النفس .

وخلاصة فكرة مفهوم مؤشر الثبات (أو ما يشار اليه خطأ بالصدق الذاتى) أن مربع معامل للصدق (أى مربع معامل الارتباط بين الاختبار والمحك) يجب أن يكون أقل من معامل الثبات ، وذلك لأن معامل الثبات يساوى مربع معامل الارتباط بين الاختبار ومتوسط الأصل الاحصائى العام الذى تشتق منه العينة . ومعنى ذلك أن مربع معامل الصدق يجب أن يكون أقل من مربع معامل ارتباط الاختبار بمتوسط الأصل الاحصائى ، وبعبارة أخرى ان الاخبار لا يمكن أن يرتبط بمتغير خارجى (وهو المحك) اعلى من ارتباطه بمتوسط أصله الاحصائى العام .

لتوضيح هذه الفكرة نعطي المثال التالي : نفرض أن معامل ثبات أحد اختبارات الذكاء بلغ ٨٧ ، معنى ذلك أننا لو استبعدنا أخطاء العينة ، فإن هذا الاختبار لا يمكن أن يتنبأ بالتحصيل المدرسي (وهو المحك) بمربع معامل صدق يزيد عن هذا المقدار (اى ٨٧) ، وعلى ذلك فإن معامل الصدق لا يجب أن يزيد بحال من الأحوال عن الجذر التربيعى لمعامل الثبات (وهو فى هذه الحالة ٩٣) . ان هذا الجذر التربيعى ليس معامل صدق (وانما يدل على الحد الذى لا يمكن أن تتجاوزه معاملات الصدق . أنه قد يعنى أن معامل الصدق قد يكون أعلى من معامل الثبات - فى ضوء هذا الحد الأقصى - الا أن هذا الافتراض النظرى لم يتحقق تجريبيا أبدا ، وجميع البيانات الامبريقية التى يجمعها الباحثون بنزاهة تتفق مع البديهية التى دارت حولها مناقشتنا السابقة وهى أن معاملات الصدق لابد أن تكون أقل من معاملات الثبات .

الفصل السادس

تصحيح الاختبارات وتفسير الدرجات

تصحيح الاختبارات :

يعد تصحيح الاختبارات بإعطاء الفرد درجة أو تقديرا ، وتفسير هذه الدرجة هو خطوة هامة ، ولو أنه في ذاته يعد مقدمة لاتخاذ قرار عملي أو تفسير علمي عن الفرد أو مجموعة الأفراد موضوع القياس . وعلى كل فإن مشكلات تصحيح الاختبارات لا تقل أهمية وخطورة عن المشكلات الأخرى التي أشرنا لها في هذا الباب ولعل أخطر هذه المشكلة ما يتصل بموضوعية التصحيح . والموضوعية objectivity هي مشكلة العلم عامة ، ومشكلة العلوم الانسانية خاصة ، وقد تناولناها في السنوات الأخيرة عدد من الفلاسفة وعلماء النفس (٢ ، ٢ ، ٤ ، ٥) ، ودون ان ندخل في تفاصيل مناقشة شائكة حول هذه المسألة ، نشير هنا الى اننا نتناول الموضوعية في اطار محدد وهو تصحيح الاختبارات . صحيح أن المسألة تتداخل مع جميع مراحل بناء الاختبار وتطبيقه وتقنيته ، وكلها تتصل بدور الباحث أو الفاحص أو المصحح .

وبالطبع فإن المعنى المباشر للموضوعية هو تحديدها « سلبيا » ولهذا كثيرا ما نجد ما تعرف بتحرر الفاحص من العوامل الذاتية كالتحيز والتعصب والآراء المسبقة والأفكار الجاهزة التي يحاول فرضها على الوقائع ، إلا أن هذا التحديد « السلبي » ليس تعريفا مفيدا للمفهوم ، وكل ما يعطيه لنا من معنى هو أن الموضوعية مقابل الذاتية . وفي رأي كاتب هذا الفصل أن هذه المقابلة الثنائية ليست هي التناول الصحيح للمسألة ، ف لغة الأضداد ليست هي اللغة الصحيحة للعلم ، ولا لمنهجه والأصح أن نتصورها على هيئة متصل يمتد على خط مستقيم يبدأ بالموضوعية الكاملة (شبه الآلية الميكانيكية) وينتهي بالذاتية الكاملة (الاعتماد الكامل على الهوى والغرض) وبينها توجد درجات كالمسماة عند سيد عثمان (٣) « البصيرة العلمية » التي تقتاتي بالخبرة .

وعلى وجه العموم نحن فيه حاجة فى مجال الاختبارات النفسية أن تكون عمليات تطبيق الاختبار وتصحيحه وتفسير درجاته مستقلة نسبيا عن الحكم الشخصى للفاحص ، حتى ولو توافر له قدر كاف من هذه البصيرة العلمية ، ولكى نتأكد من ذلك فأننا عادة ما نلجأ الى تحديد درجة « اتفاق الملاحظات والأحكام اتفاقا مستقلا » وهذا هو التعريف الذى يقترحه كاتب هذا الفصل للموضوعية . وهو تعريف موجب يتجاوز التعريفات « السالبة » المعتادة ، بالإضافة الى أنه قابل للتناول والمعالجة . فهذا « الاتفاق المستقل » فى الملاحظات والأحكام والتقديرية يمكن تحديده بحساب معامل الارتباط بين عدد من الفاحصين فى ملاحظتهم أو حكمهم أو تقييمهم - مستقلين بعضهم عن بعض - لنفس المفحوصين ، وفى هذه الحالة تصبح « الأهواء الذاتية » للفاحصين جزءا من تباين الخطأ ، ويدخل الأمر كله فى قضية ثبات الاختبار وهو نوع من الثبات وثيق الصلة بموضوعنا الراهن .

كما يمكن تحديد موضوعية الفاحص الواحد « بطريقة مماثلة وذلك بحساب معامل الارتباط بين عدد من ملاحظاته أو أحكامه أو تقديراته أو تقييماته لنفس المفحوصين والتي تختار كعينات من سلوكه ، ومن المؤكد أن هذا المعامل إذا كان منخفضا فإنه يدل على « ذاتية المصحح » ، أما إذا كان مرتفعا فقد يكون مؤشرا على « البصيرة العلمية » لديه ، بشرط أن نتأكد من عدم تداخل ملاحظاته بعضها مع بعض ، والا كان معامل الارتباط المرتفع هذا مؤشرا زائفا على هذه البصيرة ، فقد يدل فى هذه الحالة على اتساق التحيزات أو الأهواء الخاصة وفى جميع الحالات يجب التنبيه الى أن الموضوعية مسألة درجة ، وإن كانت أكثر وضوحا فى نمط الاختبارات الذى نسميه « امتحانات المقال » . وفى هذا النوع كثيرا ما يتساءل الشخص لماذا حصل على درجة منخفضة . ومن المؤكد أن السبب الرئيسى هو أنه فى هذا النوع من الامتحانات يكون من الصعب تحديد الإجابة الجيدة وبالتالي يكون من الصعب اعطاء درجة مناسبة لاستجابة ليست صحيحة صحة كاملة .

وبالطبع لا يمكن للباحث النفسى أن يجرى دراسة للسلوك كما لا يمكن أن يتخذ قرارا عمليا إذا لم تكن معايير التصحيح على درجة من الموضوعية ، ولذلك يقترح البعض لحل هذه المشكلة أما تحديد قواعد للحكم يلتزم بها المصححون للاختبارات التى تتطلب إجابات حرة أو استخدام الأسئلة

الموضوعية التي تتطلب التعرف على الاجابة الصحيحة من بين عدة اجابات .
ويستخدم ، تصحيحها مفاتيح معدة مقدما .

تصحيح الأسئلة التي تتطلب استجابات حرة :

لا يزال القياس الفردي يستخدم أسئلة ومشكلات تتطلب درجة من الحكم
اثناء التصحيح ، ومع ذلك فمن الممكن اعداد بعض الوسائل التي تتوفر فيها
درجة من الموضوعية في هذا التصحيح وبخاصة فيما يتعلق بالخصائص
الرئيسية للسلوك . فعند تقويم سلوك معين مثل الخط ، أمكن للباحثين اعداد
نموذج للتصحيح يحتوى على عينات من الخطوط تمثل مستويات مختلفة من
الجودة وعلى المدرس أن يحدد العينة الأكثر شبهة بخط التلميذ حتى يحدد
الدرجة التي يعطيها له . وكذلك يمكن اعداد مقاييس معتمدة على تقدير
الانتاج للحكم على « نوع » بعض ألوان السلوك الصناعى أو التجارى أو
الادارى . ومع أن تحديد قواعد تصحيح بعض الاختبارات التي تحتوى على
استجابات حرة حقق النجاح فى اختبارات المهارات العملية ، إلا أن هذه
الوسائل لم تنجح نجاحا مؤكدا فى تصحيح الاختبارات اللفوية ، ومع ذلك
فإن التباين بين المصححين يمكن تقليله إذا اعددنا بروتوكولات توضح نمط
التصحيح المتفق عليه بالنسبة للاجابات النموذجية . ومن أمثلة ذلك نماذج
التصحيح الخاصة باختبار ستانفورد - بينيه للذكاء ، وبروتوكولات اختبار
رورشاخ واختبار تفهم الموضوع ، وقواعد التصحيح فى اختبارات الابتكار .

وعلى الرغم من أن بروتوكولات التصحيح قد تكون مناسبة لأغراض
القياس ، إلا أننا نؤكد بعض الحذر حول تصحيح اختبارات الاستجابات الحرة
بوجه عام . فمعظم مصادر تباين الخطأ فيها تنشأ مما يمكن تسميته « المعاملة
الشخصية » للمختبرين أو المصححين ، فقد لوحظ مثلا أن بعضهم يعطى
درجات مرتفعة بصفة مستمرة بينما يميل البعض الآخر الى اعطاء درجات
منخفضة . كما أن الدرجة التي يعطيها المختبر أو المصحح تتأثر باستعداده
العقلى وتحيزه وتخمينه وأغراضه الشخصية وأهوائه ، وفى هذه الحالة
لا مناص من أن يقوموا بتصحيح أكثر من مصحح واحد تتوافر فيهم جميعا
خصائص التدريب الجيد على بروتوكولات تصحيح هذه الاختبارات أو تقدم

لهم قواعد التصحيح واضحة بحيث يمكنهم تطبيقها دون الوقوع فى أخطاء
سوء التأويل .

ولعل من المصادر الأخرى لتباين الخطأ فى هذا النوع من الاختبارات
ما يسمى أثر أو وضع الهالة الذى يحدثه المفحوص فى الفاحص ويمكن
التقليل من هذا الأثر إذا لجأنا الى ما يسمى التصحيح « الأعمى » ، أى يقوم
المختبر أو المصحح بتقدير الدرجات للأفراد دون أن يعلم نتائج الاختبارات
السابقة التى أجريت عليهم أو بالفئة التى ينتمى اليها المفحوص فى البحث
(المجموعة التجريبية أو الضابطة) أو المستوى الذى هو عليه (العفوق أو
التخلف) وبدون هذا الاحتياط يصعب علينا أن نتقضى تأثير هذا العامل
الذاتى فى الدرجة المعطاة .

تصحيح الأسئلة التى تتطلب الاختيار من متعدد :

يتكون نموذج التصحيح لما نسميه اختبارات التعرف من قائمة
الاجابات الصحيحة (أو مفتاح التصحيح) يمكن أن يستخدمه الفاحص بل
تكاد تقول أن أى شخص يمكنه أن يستخدمه فى تصحيح هذه الاختبارات ، وقد
اخترعت عدة وسائل أكثر كفاءة فى تسجيل الاستجابات فى هذا النوع من
الاختبارات وتصحيحها ، ومنها النسخ الكربونية ومنها نجد أن الوجه الخلفى
لصفحة الأسئلة (أو ورقة الاجابة) تلتصق به وتخفيه ورقة كربونية ، ويكون
مفتاح الاجابة مطبوعا على هذا الوجه الخلفى . ويختار المفحوص الاجابة
ثم يضع العلامة التى تدل على هذا الاختبار على الوجه الظاهر لورقة الأسئلة
(أو الاجابة) ، وبالطبع فإن هذه العلامة تطبعها ورقة الكربون على الوجه
الخلفى للورقة ، وعند التصحيح يقوم المختبر أو المصحح بفض الوجه الخلفى
الذى توجد فيه مربعات مطبوعة توضح أين يجب أن تكون الاجابات الصحيحة
(أى مفتاح الاجابة) ، وما على المصحح فى هذه الحالة الا عد العلامات
الكربونية التى طبعت فى المربعات الصحيحة . وبالطبع فإن هذا الاجراء
مكلف اقتصاديا .

ومن الاجراءات الشائعة استخدام أوراق مستقلة للاجابة ، وهو اجراء
يؤدى الى اختصار التكاليف ، لأن كتيب الاختبار يمكن استخدامه فى هذه

الحالة أكثر من مرة ، كما أن الاجابات يمكن تصحيحها بسهولة بمفتاح مثقوب
أو بواسطة الآلات الحديثة .

وقد ازداد في الوقت الحاضر الاعتماد على تصحيح الاختبارات
تصحيحا أوتوماتيكيا الى حد أن برامج كاملة للتقسيم النفسى فى أوربا
والولايات المتحدة تتم بالحسابات الالكترونية . وتختلف طرق التصحيح الى
الا أنها تتشابه فى عدد من العناصر منها ضرورة وجود ورقة اجابة
بمواصفات معينة ، وأن يقوم المفحوص بتسويد مكان الاجابة بالقلم الرصاص،
وتقوم الآلة بقراءة الاجابة بواسطة نظام للشفرة الموضوعية باحدى طريقتين
اولاهما خلايا ضوئية حساسة للأسود والابيض . وثانيهما التوصيل
الكهربائى لمادة الجرافيت (التى يصنع منها الاقلام) ، وفى الطريقة الأخيرة
نجد الآلة تحتوى على ، أصابع ، مكهربة توضح علامات القلم الرصاص كما
تحتوى الآلة على عداد تسجيل العدد الكلى للعلامات الموضوعية فى اماكنها
الصحيحة ، كما أن الآلة يمكنها أن تسجل عدد الأخطاء ، كما يمكنها أن
تصحح الدرجة من اثر التخمين ، ويمكن لبعض هذه الآلات ان يصحح عدة
الاف من اوراق الاجابة تصحيحا دقيقا فى الساعة الواحدة . والحاسب
الالكترونى يمكنه أن يقرأ ورقة الاجابة مرة واحدة ثم يقارن اجابات المفحوص
بالمفتاح المختزن فى ذاكرة الكمبيوتر ، وتطبع درجة المفحوص اما على ورقة
الاجابة ذاتها أو فى بطاقة خاصة ، كما يمكن للحاسب الالكترونى أن يقدم
درجات المفحوصين مبروبة (مصنفة على النحو الذى يرغب الفاحص) ،
كما يمكنه أن يقوم بتحليل هذه الدرجات مباشرة بالطرق الاحصائية الملائمة .

والصعوبة الرئيسية التى تواجهها آلات تصحيح الاختبارات ، والتى
يهتم بها المختصون فى تطبيق الاختبارات ، هى ان هذه الآلات لا يمكنها أن
تصحح تصحيحا دقيقا ما لم تكن الأماكن المخصصة للاجابة قد قام المفحوص
بتسويدها . تسويدها كاملا ونظيفا وحسب المواصفات المطلوبة والا يكون قد
وضع اجاباته فى غير مواضعها . فمن المعروف أن الآلة تعطى وحدة الدرجة
عندما تظهر العلامة التى وضعها المفحوص فى المكان الصحيح ومع ذلك فنحن
نتوقع أن تكون العلامة التى يضعها المفحوص بالقلم الخفيف أو الثقيل، أو تكون
عريضة أو ضيقة . ولذلك فلا بد أن يضع المفحوص هذه العلامة بقلم معين
وأن يملأ المساحة الكاملة المخصصة للاجابة فى ورقة الاجابة حتى نتأكد ان

الآلة سوف تعدها ولا تتجاهلها . ومن ناحية أخرى فإن الآلة قد لا تتجاهل
العلامات غير الكاملة والمحو غير النظيف للأخطاء . وقد يكون عدد الاجابات
فى ورقة الاجابة التى لا تتوفر فيها المواصفات اللازمة كثيرا ، ولذلك نجد
فى بعض المؤسسات موظفين يقومون بفحص كل ورقة قبل أن تسلم للآلة
لتصحيحها ، وهؤلاء يقومون بمراجعة اجابات المفحوصين لا بغرض
تصحيحها ولكن بغرض التأكد من أن طريقة الاجابة مناسبة للتصحيح ،
ومعنى ذلك زيادة تكاليف التصحيح . ولا شك أنه يمكن التغلب على هذه
الصعوبة بإجراء الاختبار وتطبيقه على نحو سليم .

تصحيح اثر التخمين :

من المشكلات الهامة فى تصحيح الاختبارات التى تتطلب اختيار الاجابة
من بديلين أو أكثر ما يتصل بالتخمين guessing ، وهو مصطلح عام يشير
الى أنماط السلوك التى تصدر عن المفحوص حين يستجيب لأحد البدائل
الاختيارية للاجابة على السؤال الذى لا يعرف اجابته الصحيحة . وهذه
الأنماط السلوكية عديدة ، تشمل من بين ما تشمل ما يلى (٣١ : ٥٩) .

١ - حذف بديل أو أكثر من بين البدائل الاختيارية للاجابة على السؤال
والحكم عليها بأنها خاطئة تماما ، أو حين يكون المفحوص غير متأكد من
صوابها .

٢ - الاستفادة من بعض المنبهات غير المقصودة والمتاحة بسبب صياغة
السؤال أو البدائل الاختيارية ، وقد تكون هذه المنبهات فى المعنى أو فى
البنية النحوية .

٣ - وقوع المفحوص فى الشرك الذى ينصبه له مؤلف الاختبار حين
يطرح بديلا جذابا ولكنه خاطئ .

٤ - الاستجابة على أساس ما يوجد فى أحد البدائل الاختيارية من
عناصر تجذبه دون أن يكون لديه ثقة كبيرة فى صحته .

٥ - استخدام طريقة عشوائية فى الاستجابة على نحو يعتمد كلية على

المصادفة ، وقد يصل ذلك فى تطرفه الى حد انتقاء البدائل الاختيارية التى تقع فى الاختبار فى موضع معين : مثل انتقاء الاجابة (١) او (ح) او (هـ) فى جميع الأسئلة .

ومن المؤكد أن المفحوص اذا لجأ فى اجابته الى محض التخمين معتمدا فى ذلك على العشوائية الكاملة فانه قد يحصل على بعض الدرجات ، لأن بعض تخميناته قد تقع بالمصادفة على الاجابات الصحيحة ، ولعل هذه المسألة تزداد وضوحا اذا علمنا انه فى بناء الاختبار الموضوعى الجيد الذى تقوم أسئلته على الاختيار من متعدد عادة ما تتوزع الاجابات الصحيحة على البدائل المختلفة بالتساوى ، فمثلا اذا كانت أسئلة الاختبار من نوع البدلين فان ٥٠٪ من الاجابات الصحيحة يكون رمزها (١) مثلا ، ٥٠٪ من هذه الاجابات يكون رمزها (ب) ، على أن تتوزع عشوائيا ، دون أى تتابع منظم فى الاختبار كله . فاذا كانت أسئلة الاختبار من نوع البدائل الثلاثة فان ٣٣٪ من الأسئلة يكون اجاباتها الصحيحة (١) ، ٣٣٪ رمزها (ب) ٣٣٪ رمزها (ح) على أن تتوزع مرة أخرى عشوائيا فى الاختبار ، وبالمثل اذا كانت البدائل اربعة تتوزع الاجابات الصحيحة لكل ٢٥٪ من الأسئلة على أحد الرموز (١) ، (ب) ، (ح) ، (د) ، وهكذا .

لوافترضنا أن مفحوصا اعتمد على العشوائية الكاملة فاختار الاجابات التى رمزها (١) فى جميع الأسئلة ، انه فى هذه الحالة يحصل دون حق على الدرجة ٥٠ فى الحالة الاولى ، ٣٣ فى الحالة الثانية ، ٢٥ فى الحالة الثالثة .

وحيث أننا لا نعرف فى هذه الحالة عدد التخمينات الصحيحة التى توصل اليها المفحوص بهذه الطريقة ، فأننا نستطيع استنتاج هذا العدد من عدد التخمينات الخاطئة والتى تمثلها المعادلة الآتية :

$$\frac{خ}{ن - ١} = خخ$$

حيث $خ =$ عدد التخمينات الخاطئة
 $خ =$ عدد الاجابات الخاطئة في الاختبار كله
 $ن =$ عدد البدائل الاختيارية

ويمكن استخدام المعادلة السابقة في معادلة أشمل لتصحيح اثر التخمين
والتي تستخدم في تقدير عدد الأسئلة التي يعرف الفحوص اجاباتها الصحيحة
بالفعل ويستحق عليها الدرجة الاختبارية وهذه المعادلة هي :

$$ن = ص - \left(\frac{ح}{ن - ١} \right)$$

حيث :

$ص =$ عدد الاجابات الصحيحة في الاختبار كله
 $د =$ الدرجة المصححة من اثر التخمين
 $خ =$ عدد الاجابات الخاطئة في الاختبار كله
 $ن =$ عدد البدائل الاختيارية

تفسير الدرجات

الدرجات الخام :

تعطينا معظم الاختبارات وصفا كميا مباشرا لاداء الشخص ، ويطلق على الوصف الكمي للاداء اسم الدرجة الخام raw score وقد تكون الدرجة الخام عدد الأسئلة التي أجاب عليها المفحوص في إجابة صحيحة أو خاطئة أو الزمن الذي استغرقه في الأداء أو النسبة المئوية لعدد الاجابات الصحيحة أو الخاطئة على العدد الكلي لأسئلة الاختبار .

ويفسر معظم الأفراد الدرجة الخام ويستخدمونها دون معرفة لجوانب القصور فيها لأن هذه الدرجات تتميز بالسهولة في الحصول عليها وبالفقنا لها من خبرتنا الطويلة بالامتحانات المدرسية . ومع ذلك فإننا نستطيع القول أن الدرجة الخام في الاختبار النفسي لا معنى لها في ذاتها وليست لها أية دلالة ، ولا يمكن أن تفسر إلا باللجوء الى أساس آخر للمقارنة ، ولعلنا نعود بهذا الى التعريف الاساسي للاختبار النفسي . ومن المهم أن نؤكد في هذا الصدد أننا لا نستطيع تفسير الدرجات في الاختبار النفسي كما نفعل بالقيم التي نحصل عليها من مقياس الظواهر الطبيعية (الطول مثلا) .

فمقاييس الظواهر الطبيعية هي من نوع مقاييس النسبة ، أي لها صفر مطلق وتتألف من وحدات متساوية ، ولذلك فإن هذه المقاييس تسمح لنا بتفسير مباشر للقيم التي نحصل عليها منها . فنقول مثلا أن طول الطفل (س) نصف طول والده (ص) باستخدام درجات المتر مباشرة . أما في القياس النفسي فإننا لا نستطيع مثل هذا القول لأن المقاييس في هذه الحالة لا تكون في أحسن حالاتها إلا من نوع مقاييس المسافة . لنفرض مثلا أن الطفل (س) حصل على درجة مقدارها ١٠٪ من أسئلة اختبار ، في معاني الكلمات فهل يعني ذلك أن ذلك الطفل لا يعرف الا عشر الكلمات التي يجب أن يعرفها ؟

الاجابة على ذلك بالنفي لأن الفاحص قد لا يكون مهتما الا بالكلمات الصعبة غير المألوفة للطفل ولهذا نجد أن محصوله اللفظي قد يكون أكبر في اختبار أقل صعوبة . وحتى لو حصل الطفل على الدرجة صفر في هذا الاختبار فإن ذلك يعني أن الطفل ليست لديه قدرة لغوية على الإطلاق . فالفرق بين (س) الذي حصل على الدرجة صفر في الاختبار والطفل ص الذي يحصل على النهاية القصوى أي يجيب على أسئلة الاختبار جميعا ، هو فرق في القدرة على

معرفة عينة من الكلمات يتكون منها الاختبار اشتقت من بين الالف الكلمات ، فالدرجة التى تدل على ٦٥٪ صواب فى أحد الاختبارات قد تتساوى مع الدرجة التى تساوى ٣٠٪ فى اختبار آخر أصعب ، ومع ٨٠٪ فى اختبار ثالث أسهل . فمستوى صعوبة الكلمات التى يتألف منها الاختبار تحدد الدرجة الخام للمفوحص .

معنى ذلك أن الفرق بين الأفراد فى الدرجات الخام للاختبارات النفسية لا يدل بالضرورة على وجود مسافات « حقيقية » بين الأفراد لأن هذه الاختبارات لا يتحقق فيها الشرطان الأساسيان لمقاييس المسافة وهما الصفر المطلق والوحدات المتساوية . فمثلا فى اختبار الذكاء قد يحصل التلميذ (١) على الدرجة ٥٣ والتلميذ (ب) على الدرجة ٥٦ والتلميذ (ح) على الدرجة ٥٩ ، أى تكون الفروق بين الدرجات الخام متساوية ، ومع ذلك لا يدل هذا على أن الفرق بين (ح) و (ب) يساوى تماما الفرق بين (ب) و (١) ، ولا يمكن الاجابة على مثل هذا السؤال اجابة مؤكدة ، لأن الفروق بين الدرجات انما تعتمد على طبيعة الأسئلة (أو الوحدات) التى تكون الاختبار كما قلنا . ولهذا لابد من أن نلجأ الى بعض الأسس التى تتجاوز الدرجات الخام فى المقارنة بين المفحوصين وتفسير درجاتهم والحكم على أدائهم .

ويضيف فؤاد أبو حطب (٩) هذه الأسس التى يجب استخدامها فى المقارنة ثم التفسير والحكم على الدرجة الخام الى ثلاثة أنواع هى : المعايير norms ، والمحككات Criteria ، والمستويات Standards وفيما يلى تفصيل الأسس الثلاثة :

(أولا) المعايير

كانت « المعايير » أول ما ظهر فى ميدان القياس النفسى للتغلب على الصعوبات المتضمنة فى تفسير الدرجات الخام والحكم على أداء المفحوص وتعد المعايير أساسا لتفسير أداء المفحوصين والمقارنة بينهم فى ضوء أدائهم الفعلى ، وتتحدد فى ضوء الخصائص الواقعية لهذا الأداء ، وتعتمد فى جوهرها على الأداء الاختبارى لعينة ممثلة للأصل الاحصائى موضع الاهتمام

تسمى عينة التقنين Standardization sample وهكذا تتحدد المعايير تجريبيا بما تستطيع مجموعة ممثلة من الأفراد أداءه (عينة ممثلة) ، ثم نشير الى الدرجة الخام التي يحصل عليها المفحوص في ضوء توزيع الدرجات التي تحصل عليها هذه العينة ليتحدد موقعه ومكانته في هذا التوزيع : هل تتفق درجته مثلا مع الأداء المتوسط لعينة التقنين ؟ هل تقع في مستوى أقل قليلا (او أعلى قليلا من المتوسط ؟ هل تقع في الطرف الأعلى من التوزيع ؟ وهكذا .

وحتى يمكن أن نحدد على وجه الدقة موضع الفرد بالنسبة لعينة التقنين فإن الاختبار النفسى يتحول الى نوع من مقاييس الرتبة أو مقاييس المسافة وتتحول الدرجة الخام الى هذا النوع من القياس عن طريق البحث عن مقابلها من رتبة أو مسافة وتسمى هذه القيم المقابلة للدرجات الخام بالدرجات المشتقة derived scores . وهذه الدرجات المشتقة تفيد في تحقيق غرضين : أولهما تحديد الوضع النسبى للفرد فى العينة المعيارية وتقويم أدائه فى ضوء أداء الآخرين المماثلين له أو المختلفين عنه) ، وثانيهما أنها تعطينا مقاييس قابلة للمقارنة سواء بين الأفراد أو بين الاختبارات .

الدرجات المشتقة

وحتى عهد قريب كانت المعايير هى الأساس الوحيد الذى يتطلبه علماء القياس النفسى للمقارنة بين الدرجات الخام وتفسيرها ، حتى أنها كانت تعد شرطا جوهريا لجودة الاختبارات النفسية على اختلاف أنواعها وأغراض استخدامها . الا أنه مع التطور المعاصر فى النظرية الأساسية للاختبارات النفسية وظهور أنواع جديدة منها وأغراض متنوعة لها ، وخاصة ما يسمى الاختبارات المرجعة الى المحك توجه الباحثون الى اعتبار المعايير أساسا واحدا من بين عدة أسس لهذه المقارنة ، واقتصر دورها على نوع واحد فقط من هذه الاختبارات هو ما يسمى « الاختبارات المرجعة الى المعيار » ، ونعرض فيما يلى المعايير الشائعة الاستخدام فى هذا النوع من الاختبارات .

المعايير الارتقائية او النمائية :

تشير المعايير الارتقائية و النمائية developmental الى المعنى الذى يعطى للدرجة الخام ويدل على اقتراب أو ابتعاد نمو المفحوص عن

المسار الطبيعي لنمو العملية النفسية موضع القياس . وعلى الرغم من شيوع هذا النوع من المعايير (وخاصة في صورة معيار العمر العقلي بالنسبة لاختبارات الذكاء) إلا أنه ليس دقيقاً ولا يصلح للمعالجة الإحصائية المحكمة ويجب ألا تتجاوز استخدامه حدود الأغراض الوصفية في مجال البحث أو المواقف الكليникаية .

٦ - معايير العمر والفرق الدراسية : لقد كان العالم الفرنسي ألفريد بينيه Binet أول من استخدم مفهوم العمر العقلي mental age في تعديله لمقياسه المشهور في الذكاء عام ١٩٠٨ . وفي مقاييس العمر كما في عينة التقنين ، ومتطابق الدرجة العمرية التي يحصل عليها الفرد مع أعلى مستوى عمري يمكن أن يصل إليه بنجاح . فإذا استطاع طفل عمره الزمني ١٠ سنوات أن يجيب بنجاح على أسئلة ١٢ سنة فإن عمره العقلي يصبح ١٢ سنة ، ومعنى ذلك أنه يسبق عمره الزمني بعامين . لأن أدائه يساوي أداء الطفل المتوسط من عمر زمني أكبر منه بعامين .

وقد لوحظ أن أداء المفحوص في الاختبارات التي تعتمد على معيار العمر العقلي يظهر قدراً من التشتت . وبعبارة أخرى فإن الطفل قد يفشل في أداء بعض الأسئلة التي تقع في مستوى عمري أدنى من عمره العقلي ، كما يمكنه الإجابة على بعض الأسئلة التي تقع في مستوى أعلى منه . ولهذا السبب نلجأ عادة إلى حساب العمر القاعدي basal age ويقصد به أدنى مستوى عمري يستطيع الطفل أن يجيب على جميع أسئلته والأسئلة السابقة عليه ، ثم تضاف إليه درجات جزئية - في صورة كسور من السنة أو شهر - تدل على جميع الأسئلة التي أجاب عليها من مستوى أعلى من العمر القاعدي وبذلك يصبح العمر العقلي للطفل هو مجموع العمر القاعدي والأشهر الإضافية التي تدل على ما حصل عليه الطفل من مستويات عمرية أعلى منه حتى يصل إلى سقف الاختبار والذي يدل على المستوى العمري الذي يعجز الطفل عن الإجابة على جميع أسئلته والأسئلة اللاحقة عليه .

وقد نلجأ إلى استخدام معايير العمر العقلي في الاختبارات التي لا تنقسم مفرداتها إلى مستويات عمرية كما هو الحال في اختبار ستانفورد - بينيه . وفي هذه الحالة تكون الدرجة الخام هي عدد الإجابات الصحيحة

أو الخاطئة على جميع أسئلة الاختبار - أو الزمن الكلى الذى استغرقه
المفحوص فى حله وتحسب متوسطات الدرجات الخام التى تحصل عليها
الجموعات العمرية المختلفة فى عينة التفتين ، وتكون تلك المتوسطات هى
معايير الأعمار العقلية للاختبار كما تستخدم معايير عمرية أخرى مثل العمر
الطولى والعمر الوزنى وعمر التصنين والعمر الرسفى وعمر قبضة اليد وعمر
القراءة وغير ذلك من الخصائص النفسية والجسمية ذات الصبغة الارتقائية
ويمكن بالطبع استخدام هذا الاجراء للحصول على معايير الفرق الدراسية
فى حالة الاختبارات التحصيلية المقتنة حيث تستخدم فئات الفرق الدراسية
بدلا من فئات الأعمار .

ومن أهم مشكلات وحدة العمر العقلى ووحدة الفرق الدراسية
كمعيار اختبارى أنهما لا يتقيمان بمعدل ثابت فى مختلف مراحل العمر أو
مراحل الدراسة ، فالنمو العقلى أسرع وبالتالي يكون مقداره أكبر فى
الطفولة المبكرة منه فى الطفولة المتأخرة ، ويزداد بطئا وبالتالي يقل مقداره
بإقتراب الشخص من حد النضج ، كما أن النمو التحصيلى ليس متكافئا
فى جميع الفرق الدراسية ومعنى ذلك أن مقدار النمو الذى تتضمنه هذه
الوحدات - بلغة القياس النفسى - يجعلها غير متساوية من عمر لآخر ،
وبالتالى يجعلها غير قابلة للمقارنة .

وقد تتضح هذه الفكرة إذا أعطينا مثالا من العمر الطولى ، اننا فى
هذه الحالة نجد أن الفرق بالمستقيمات بين العمر الطولى للطفل فى سن
٢ أو ٤ أكبر من الفرق بين العمر الطولى فى سن ١٠ ، ١١ ، ومعنى ذلك
أن عاما واحدا من « التفوق » أو « التخلف » فى سن ٤ سنوات يمثل اختلافا
عن المستوى المتوسط أكبر كثيرا من مقدار يساويه من التفوق أو التخلف فى
سن ١١ سنة .

٢ - نسبة الذكاء التقليدية : ابتكر علماء النفس معيار نسبة الذكاء
intelligence quotient أو I.Q. للحصول على تفسير موحد للذكاء
بصرف النظر عن العمر الزمنى للمفحوص . وقد أشار الى أهمية ما يسمى
السنة العقلية شيرن وتولمان . إلا أن هذا المعيار استخدم لأول مرة فى

مقياس ستانفورد - بينيه الذى أعده ترمان - ام ١٩١٦ ويمكن الحصول على نسبة الذكاء كما يلى :

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠$$

فاذا تعادل العمر العقلى والعمر الزمنى تكون نسبة الذكاء فى هذه الحالة ١٠٠ وتدل على أداء متوسط أو عادى ، وإذا قلت عن ١٠٠ دلت على تخلف أو ضعف ، وإذا زادت عن ١٠٠ دلت على التفوق .

وتفيد نسبة الذكاء فى ضبط الضمور العقلى لطفل عمره الزمنى ٤ سنوات هو ٣ تصبح نسبة ذكائه فى هذه الحالة ٧٥ فاذا بلغ هذا الطفل سن ١٢ وكان عمره العقلى ٩ تظل نسبة ذكائه ثابتة أى ٧٥ . أى أن نسبة الذكاء تدل على نفس الوضع النسبى للفرد فى الجماعة سواء فى سن ٤ سنوات أو فى سن ١٢ سنة . وبعبارة أخرى فإن نسبة الذكاء يمكن استعمالها فى المقارنة فى مختلف الأعمار ، بمعنى أن تفسير نسبة الذكاء يظل هو نفسه بصرف النظر عن العمر الزمنى للمفحوص .

ولابد أن نشير فى هذا الصدد الى أنه لكى تظل نسبة الذكاء ثابتة فإن ذلك لا يمكن - نظريا - إلا فى حالة واحدة فقط وهى أن وحدة العمر العقلى تتناقص ، (أى تصبح أصغر) تناقصا يتناسب تناسباً مباشراً مع العمر الزمنى . وهذا شرط ضرورى حتى يمكن استخدام نسبة الذكاء استخداماً له معنى .

وحين تتناقص وحدة العمر العقلى تبعاً للتقدم فى العمر على هذا النحو فإن الفروق الفردية التى تقاس بهذه الوحدة يفترض فيها بالطبع أنها تتزايد بنفس النسبة . ويشبه ذلك من الناحية الحسابية أن الفروق الفردية فى الطول تكون أكبر ١٢ مرة حينما تقاس بالبوصات عنها لو قيست بالأقدام وذلك لأن البوصة هى $\frac{1}{12}$ من القدم وبنفس المنطق فإن الفروق الفردية أو التباين فى العمر العقلى يجب أن تكون فى سن ١٤ ضعفها سن ٧ سنوات ، ولا يمكن لنسب الذكاء أن تظل ثابتة إلا إذا تزايد الانحراف المعيارى للأعمار

العقلية بنفس نسبة زيادة الأعمار الزمنية ، وبهذا الشرط وحده يمكن أن يكون لنسبة الذكاء نفس المعنى فى مختلف الأعمار .

وقد أمكن تحقيق هذا الشرط فى الطبقات الحديثة من اختبار ستانفورد - بينيه فقد اتضح من دراسة انتشار الأعمار العقلية من سن ٦ حتى سن ١٨ أن هناك اتجاهها نحو زيادة التباين فى الأعمار العقلية تبعاً لزيادة الأعمار الزمنية . ولكننا نجد فى كثير من الأحيان عدم تساوى الانحرافات المعيارية لنسب الذكاء فى مختلف الأعمار ، وهذه الصعوبة أمكن التغلب عليها فى اختبار ستانفورد - بينيه بأعداد جدول تصحيح يستخدم فى تصحيح نسب الذكاء فى مستويات عمرية معينة .

وهناك مشكلة أخرى فى تطبيق معيار نسبة الذكاء نجدها فى اختبار البالغين والكبار فالواقع أن مفهوم معايير العمر على إطلاقه يدل دلالة مباشرة على أن فائدته مقصورة على الأطفال لما يتضمنه من معانى ارتقائية . أما بالنسبة للراشدين والكبار فإن ظواهر النمو ليست على نفس الدرجة من الوضوح إلا فى حالات تدهور الشيخوخة ولذا نجد أن فى اختبار مثل ستانفورد - بينيه لا يتحسن أداء البالغ المتوسط كثيراً . صحيح أنه قد أضيف إلى الاختبار مستويات رفيعة لذكاء الكبار حتى يتوافر فيه سقف مناسب وبالتالي يمكننا الحصول على أعمار عقلية أعلى ، ولكننا إذا وجدنا شخصاً معيناً يحصل على عمر عقلى مقداره ٢٠ فى هذا الاختبار فإن ذلك لا يسمح لنا بتفسير مؤكد وعلى نفس الدرجة من الوضوح كما هو الحال فى الأعمار العقلية الدنيا . فمن المؤكد أن العمر العقلى ٢٠ لا يمكن تحديده بما يمكن أن يؤديه الشخص المتوسط البالغ من العمر ٢٠ سنة ، فى هذا الاختبار ، أما إذا اخترنا شخصاً بالغا ضعيف العقل يقل عمره العقلى عن حدود النضج فإن استخدام مفهوم العمر العقلى فى هذه الحالة عنه يتشابه باستخدامه مع الأطفال . أما فى حالة الراشدين والكبار من الأسوياء أو المتفوقين فلا مناص من استخدام أنواع أخرى من المعايير ليس من بينها معيار العمر لصعوبة تحديدهم عندهم .

أما المشكلة الثالثة التى تواجهنا حينما نفسر الدرجات الخام بدرجات العمر تتمثل فى أن هذا النوع من المعايير لا يستخدم إلا فى الوظائف النفسية

والعمليات العقلية التي تتغير تغييرا واضحا ومتسقا مع النمو . أما بالنسبة للسمات التي لها علاقة ضئيلة بالعمر الزمني أو لا تكون لها علاقة مباشرة به لا يجوز بالطبع استخدام معيار العمر معها ، ومن أمثلة ذلك مقاييس سمات الشخصية ، فبعض هذه السمات يظل ثابتا في معظم مراحل نمو الفرد .

ولأسباب تعود الى هذه المشكلات وغيرها توقف علماء القياس النفسي المعاصرون عن استخدام نسبة الذكاء على النحو المشار اليه ، وأصبح يطلق عليها اسم نسبة الذكاء التقليدية تمييزا لها عن أنواع أخرى من المعايير اكبر دقة سنشير اليها فيما بعد ولعل في هذا ما ينبه الباحثين في مصر والبلاد العربية الى أن يتوجهوا في بناء الاختبارات النفسية وفي استخداماتها في الأغراض المختلفة (وخاصة الأغراض الكلينيكية) الى المعايير الأخرى المناسبة غير نسبة الذكاء التقليدية .

٣ - المئينيات : يعبر عن الدرجات المئينية percentiles في صورة نسب مئوية للأفراد من عينة التقنين الذين يقعون أدنى من درجة خام معينة . وتدل على الوضع النسبي للفرد بمقارنته بعينة التقنين هذه . وهي تعد من نوع الرتب فيما عدا أننا في الترتيب نبدأ عادة باعتبار الشخص الذي يقع في القمة - أي أفضل أفراد المجموعة من حيث الأداء - في الرتبة الأولى ، بينما في المئينيات فإننا نبدأ العد من أسفل ، وبالتالي فإن أدنى المئينيات يدل على أسوأ مكانة للفرد .

ويقابل المئيني الخمسون منتصف جماعة التقنين أو وسيطها فإذا زاد المئين عن ٥٠ يدل هذا على أداء أعلى من المتوسط ، وإذا قل عنه دل على أداء أقل من المتوسط . ويعرف المئيني ٢٥ ، والمئيني ٧٥ احصائيا بالربيع الأدنى والربيع الأعلى من التوزيع ، وهما مثل الوسيط يعطيانا معالم تصف توزيع الدرجات ومقارنته بغيره من التوزيعات ، وهي كالوسيط والربيع في طريقة حسابها ، كما أنها تقسم التوزيع التكراري الى نقاط ، وهي في حالة المئينيات تصل الى ٩٩ نقطة .

ويجب الا نخلط بين المئينيات ودرجات النسب المئوية الشائعة ، فدرجات النسب المئوية هي نوع من الدرجات الخام يعبر عنها في ضوء نسب مئوية

للإجابات الصحيحة . أما المئينيات فهي درجات مشتقة يعبر عنها في ضوء نسب مئوية للمفحوصين فإذا افترضنا أن لدينا ٤٠ شخصا منهم ٢٧ حصلوا على درجات أعلى من الشخص (١) ، ١٢ حصلوا على درجات أقل منه ، فأنشأ في هذه الحالة نضع هذا الشخص بين المجموعتين معتبرين أنه تفوق على ٢١٪ من العدد الكلي للمجموعة ، وتكون درجته المئينية . والدرجة الخام التي تقل عن أي درجة حصل عليها أفراد عينة التقنين تكون رتبها المئينية صفرا ، والدرجة الخام التي تزيد على أعلى درجة حصل عليها أفراد عينة التقنين تكون رتبها المئينية ١٠٠ وهذه المئينيات لا تتضمن بحال من الأحوال أن تكون الدرجات الخام المقابلة لها مقدارها صفرا أو تدل على الأداء الكامل أو الحد الأقصى للدرجة الخام .

وللمئينيات ميزات عديدة : فهي سهلة الحساب ميسورة الفهم ، كما أن استخدامها يصلح لكل أنواع الاختبارات ، كما تصلح للأطفال والمراهقين والراشدين على حد سواء . إلا أن أهم مشكلاتها تنشأ من عدم تساوي وحداتها وخاصة عند أطراف التوزيع ، فإذا كان توزيع الدرجات الخام يقترب من التوزيع الاعتدالي فإن الفروق بين الدرجات الخام الأقرب إلى الوسيط أو مركز التوزيع تتزايد (أي تنتشر الحالات الكثيرة التي تقع في المنتصف وتتقنت إلى مئينيات كثيرة) بينما تتضاءل هذه الفروق عند الأطراف (أي تتجمع الحالات القليلة التي تقع عند الأطراف بعضها إلى بعض في عدد قليل من المئينيات) . فإذا أعدنا رسم التوزيع التكراري بيانيا باستخدام المئينيات نحصل على ما يسمى التوزيع المستطيل .

ولا بد أن نشير إلى أننا لا نستطيع استخراج متوسطات الدرجات المئينية ، لأن هذا المتوسط يختلف في مقداره عما لو استخرجنا المتوسط الحسابي للدرجات الخام ثم حولناه إلى درجة مئينية . وبدلاً من اللجوء إلى المتوسط الحسابي لقياس النزعة المركزية للدرجات المئينية يمكن للباحث أن يعتمد على الوسيط باعتبارها المقياس الإحصائي المناسب للنزعة المركزية في هذه الحالة .

كما يجب الإشارة إلى أن الدرجات المئينية لا يمكن المقارنة بينها ما لم تكن هذه المقارنة ممكنة من الجماعات التي تشتق منها فإذا اختلفت

الجماعات ، كأن تكون أحداها مثلا من تلاميذ مدرسة ثانوية صناعية والثانية من عمال تلمذة صناعية والثالثة من طلاب كلية للهندسة فان الدرجات الخام التي تقابل مئيني معين في كل حالة من هذه الحالات تختلف ، كما ان الدرجة الخام الواحدة في كل مجموعة من هذه المجموعات الثلاث تقابل مئينيات مختلفة ومن ذلك مثلا ان الدرجة الخام التي تقابل الدرجة المئينية الـ ٦٠ مثلا لطلاب الجامعة قد تقابل مثلا المئيني الـ ٧٥ في مجموعة طلاب المدرسة الثانوية الصناعية وتقابل المئيني الـ ٩٠ في مجموعة عمال التلمذة الصناعية .

بعبارة أخرى اذا لجأنا الى استخدام الدرجات المئينية لابد للباحث ان يضع في اعتباره عينة التقنين التي تشتق منها هذه المعايير ، ولو ان ذلك لا يحول دون استخدام مجموعة من جداول التحويل الى درجات مئينية تسمح للمختبر أو الباحث أن يقارن الشخص الذي يختبره بمجموعات مرجعية مختلفة .

٤ - الدرجات المعيارية : يتزايد استخدام الدرجات المعيارية Satndard scores كمعايير في الاختبارات الحديثة ، وتعتبر هذه المعايير أفضل صورة لتحويل الدرجات الخام لأنها تعبر عن بعد الدرجة الخام التي يحصل عليها المفحوص عن المتوسط منسوباً الى الانحراف المعياري للتوزيع ، أي أن :

$$\frac{\text{الدرجة الخام} - \text{المتوسط}}{\text{الانحراف المعياري}} = \text{الدرجة المعيارية}$$

$$\frac{س - م}{ع} = \text{أو } ز$$

ويطلق على هذا النوع من الدرجات المعيارية الخطية Linear standard scores ويقصد بذلك أننا نحفظ بالعلاقات العددية الدقيقة للدرجات الخام الأصلية ، فنحن نطرح من الدرجة مقدارا ثابتا (المتوسط) ، ونقسمها على مقدار ثابت (الانحراف المعياري) . ومعنى ذلك أن السعة النسبية للفروق بين الدرجات

المعيارية التي تشتق من هذا التحويل الخطى تتطابق تماما مع الفروق بين الدرجات الخام ، كما تتكرر جميع خصائص التوزيع الأصلي للدرجات الخام في توزيع هذه الدرجات المعيارية .

ولعلنا نذكر ان من أهم أسباب تحويل الدرجات الخام الى درجات مشتقة من المعايير عامة هو جعل الدرجات التي نحصل عليها من مختلف الاختبارات قابلة للمقارنة ، وهذا الفرض لا يمكن ان يتحقق للدرجات المعيارية الخطية الا اذا كانت التوزيعات التكرارية للاختبارات المختلفة متشابهة ، فاذا كان أحد التوزيعين اعتداليا والاخر ملتويا فلا بد في هذه الحالة من اجراء تحويلات غير خطية للدرجات الخام حتى تتواءم مع نمط معين من انماط التوزيع ، وعادة ما يختار الباحثون في مجال الاختبارات المرجعة الى المعيار التوزيع الاعتدالي لأسباب نظرية خالصة ، ومعنى ذلك ان التوزيعات المختلفة يجب ان تتحول في هذه الحالة الى توزيعات تتفق مع خصائص المنحنى الاعتدالي ، ثم تحول الدرجات الى درجات معيارية اعتدالية باستخدام جداول مساحات وارتفاعات المنحنى الاعتدالي *normalized standard scores* وتصبح لهذه الدرجات المعيارية الاعتدالية في هذه الحالة نفس خصائص الدرجات المعيارية الخطية (أى متوسطها صفر وانحرافها المعيارى واحد صحيح) .

الا ان ما يجب ان ننبه اليه هو ان الدرجات المعيارية الخطية لا تتطابق مع الدرجات المعيارية الاعتدالية الا في حالة واحدة فقط وهي ان تكون التوزيعات التكرارية التي حسبت منها الدرجات المعيارية الخطية اعتدالية بالطبع . كما ننبه ايضا الى ان التحويلات الاعتدالية لا يجب ان تتم الا اذا توافرت شروط الاعتدالية من حيث السمة المقيسة والعينة المختبرة والمقياس المستخدم ، وهي شروط فصلها فؤاد أبو حطب في موضع آخر (٩) .

ومن الواضح ان حساب الدرجات المعيارية الخطية يتضمن ان تكون هذه الدرجات سالبة او موجبة . فالدرجات السالبة تدل على أداء أدنى من المتوسط ، والدرجات الموجبة تدل على أداء أعلى منه ، والدرجة الخام التي تساوى المتوسط تقابل الدرجة المعيارية صفرا . كما أننا نلاحظ على هذه

الدرجات المعيارية أنها تتضمن كسورا عشرية حتى تتوافر لدينا معلومات كافية نميز بها بين مختلف المفحوصين وخاصة اذا علمنا ان المدى الكلى لأغلب المجموعات لا يتعدى نطاق ٣ انحرافات معيارية .

وللتغلب على هاتين الصعوبتين اقترحت عدة تعديلات تهدف الى تحويل الدرجات المعيارية الى صورة اكثر ملاءمة وأيسر في الفهم . وتعتمد هذه التعديلات على بعض الخصائص الرياضية للدرجات المعيارية . والتي تتمثل في أن متوسط هذه الدرجات يساوى صفرا ، وانحرافها المعياري يساوى الواحد الصحيح ، ولذلك نستطيع أن ندخل عليها ما نشاء من تعديلات باستخدام متوسطات وانحرافات معيارية جديدة ، وفي تحويل الدرجات المعيارية الأصلية الى المقياس الجديد يتطلب الأمر ببساطة أن نضربها في الانحراف المعياري الجديد ونجمعها الى المتوسط الجديد . أي أن الدرجة المعيارية المعدلة تحسب كما يلي :

الدرجة المعيارية المعدلة = الدرجة المعيارية الأصلية × الانحراف المعياري الجديد + المتوسط الجديد .

$$D' = (D \times C) + M$$

وقد كثرت الدرجات المعيارية المعدلة وتعددت ، ويلخص الجدول رقم (٦) بعض أنواعها الشائعة ، وقد يكون أكثرها شيوعا في الوقت الحاضر ، وبخاصة في اختبارات الذكاء ، نسبة الذكاء الانحرافية ، وهي تختلف تماما عن نسبة الذكاء التقليدية التي عرضناها انفا ، فنسبة الذكاء الانحرافية درجة معيارية معدلة متوسطها الجديد ١٠٠ وانحرافها المعياري الجديد قد يكون ١٥ (كما هو الحال في الطبقات الحديثة من مقياس ستانفورد - بينيه) أو ١٦ (كما هو الحال في اختبارات وكسلر) .

وقد أثر بناء اختبارات الذكاء بمفهوم نسبة الذكاء في معيارهم القائم على فكرة الدرجة المعيارية (نسبة الذكاء الانحرافية) لأسباب تاريخية ترجع في جوهرها لاستخدام المصطلح لفترة طويلة في علم النفس ، واختير الانحراف المعياري الجديد (١٦ أو ١٥) في ضوء التحليل الاحصائي

للانحرافات المعيارية لنسب الذكاء التقليدية والتي بلغ وسيطها في مقياس ستانفورد - بينيه طبعة ١٩٢٧ حوالى ١٦ ، أما المتوسط ١٠٠ فقد اختير لأنه يشير في نسبة الذكاء التقليدية الى المتوسط ٠ وهكذا تتوافر في نسبة الذكاء الانحرافية (كدرجة معيارية معدلة) جميع الخصائص « الشكلية » ، لنسبة الذكاء التقليدية الا انها تختلف عنها تماما في الأسس الرياضية وطرق الحساب كما بينا بالاضافة الى اختلافهما في المعنى التفسيري فنسبة الذكاء التقليدية هي نسبة العمر العقلى الى العمر الزمنى ، أما نسبة الذكاء الانحرافية فتدل على اقتراب أو ابتعاد عن المتوسط بمسافات من الانحراف المعيارى ٠

جدول رقم (٦) بعض أنواع الدرجات المعيارية المعدلة

المتوسط	الانحراف المعيارى	الدرجة المعيارية التى تقابل انحرافا معياريا واحدا أعلى من المتوسط	الدرجة المعيارية التى تقابل انحرافين معيارين أقل من المتوسط	اسم الدرجة المعيارية
صفر	١	١	- ٢	الدرجة المعيارية الأصلية
٥	٢	٧	١	الدرجة التساعية أو التساعى المعيارى
٥٠	١٠	٦٠	٢٠	الدرجة الثائية
١٠٠	١٦ و ١١٥	١١٥ أو ١١٦	٧٠ أو ٦٨	نسبة الذكاء الانحرافية

تعليق على المعايير : المعايير اذن هي أسس للحكم على أداء المفحوصين والمقارنة بينهم في ضوء أدائهم الفعلى في الاختبارات ولأنها ذات طبيعة كمية يمكن استخدام الحاسبات الالكترونية معها بحيث يقدم لنا الكمبيوتر تفسيراً للدرجات الخام ٠ وحتى يمكن المقارنة بين المعايير الخاصة بالاختبارات المختلفة ، او بين المعيار الواحد المستخدم مع المفحوص في مراحل مختلفة من تطبيق الاختبار ، لابد من أن يكون الاختبار المستخدم والذي تشتق من درجات عينة تقنيه المعايير الخاصة به معلوما للمفحص ٠ فالمعايير لا تستخدم لأغراض المقارنة المجردة المطلقة ٠

وتفيدنا معرفة طبيعة الاختبار المستخدم في القياس في تحديد محتواه

وما يقيسه ، ونود أن ننبه هنا الى الاختبارات التي تبدو لنا متشابهة فيما تقيس (كاختبارات الذكاء) هي في الواقع مقاييس لعمليات نفسية مختلفة .

نفسه

وحتى يدرك الفاحص مغزى المعايير ومعناها عليه أيضا أن يتحقق من القيم العددية المستخدمة في حساب المعايير . فحتى يستطيع المقارنة بين نسب الذكاء الانحرافية لاختبارين للذكاء لابد أن تكون واضحة تماما وحدات القياس المستخدمة في حساب هذا المعيار - فنسب الذكاء الانحرافية (المتوسط جديد مقداره ١٠٠) تختلف تبعاً للقيم المختارة للانحراف المعياري الجديد (١٥ او ١٦ مثلا) .

ولكى يفهم الفاحص طبيعة المعيار ويستخدمه استخداما صحيحا لابد أن يتعرف بعناية على عينة التقنية المستخدمة في حساب معايير الاختبار ، وعلينا دائما أن نعلم أن معايير أى اختبار لا يمكن أن تتجاوز حدود خصائص عينة تقنيته بشرط أن تتوافر فيها دائما شروط العينة الجديدة وخاصة شرط التمثيل ، فإذا أردنا استخدامه مع مفحوصين ذوي خصائص مختلفة لابد من إعادة تقنيته على عينات جديدة . وهنا تجدنا ازاء موقف هام بالنسبة للمعايير ، وهو أن الاختبار الواحد قد تكون له معايير متعددة بعضها قد يكون عاما على المستوى القومي او اقل عمومية على المستوى المحلي ، بل قد يصل الأمر الى حساب معايير خاصة لعينات تقنين مختلفة تشتق من مختلف الفئات التي يصنف اليها المجتمع الواحد لو الثقافة الواحدة ، وفي هذه الحالة ينسب المفحوص الى الأصل الإحصائي الذي ينتسب اليه ويكون الحكم على أدائه وتفسيره أقرب الى العدالة .

صحيح أنه ظهرت في السنوات الأخيرة بعض الاتجاهات العملية والنظرية تعين على بناء مقاييس موحدة « متحررة من أثر العينة » وتصلح للتطبيق على الأفراد والجماعات من مستويات مختلفة في القدرة ولقياس محتوى من مستويات مختلفة من الصعوبة ، ومن هذه الطرق ما يسمى استخدام المجموعة المرجعية الثابتة *fixed reference group* ونماذج السمة الكامنة *latent trait models* ، وهي تعتمد على مجموعة من الأسئلة أو الاختبارات الفرعية تسمى أسئلة أو اختبارات الارتكاز *anchor*

وتختار بموصفات معينة بحيث تتجاوز العينات والاختبارات والأسئلة ، الا
أن هذه الطرق جميعا تهتم بأحد مصادر الفروق بين الأفراد والجماعات وهو
صعوبة الأسئلة ، وتوجد مصادر أخرى لا تقل أهمية .

وتبقى أخيرا الإشارة الى أن المعايير كأسس للحكم على الدرجات
الخام وتفسيرها تصلح للاختبارات التى تسمى فى الوقت الحاضر الاختبارات
المرجعة الى المعيار دون سواها ، ولا تصلح للانواع الأخرى التى تستخدم
معها المحكات والمستويات كما سنبين فيما بعد .

المحكات :

يمكن للدرجات الخام أن تفسر أيضا فى ضوء الأداء فى محك متوقع ،
والمحك كما أشرنا من قبل هو أساس خارجى مستقل للحكم على الأداء فى
الاختبار ، كما هو الحال فى برامج التدريب أو مقاييس الكفاية التعليمية
والمهنية ، ويتفق هذا المعنى للمحك مع استخدامه فى بحوث الصدق ، وإلى
هذا المعنى يجب أن يشير ما يسمى فى الوقت الحاضر الاختبارات المرجعة
الى المحك الا أن المعنى الاصطلاحي الشائع لهذا النوع من الاختبارات
يخرجها من المعنى الحالى للمحك ويدخلها فى فئة جديدة سنشير إليها تحت
اسم « المستوى » . وعلى ذلك فإن أسس الحكم على ما يسمى « الاختبارات
المرجعة الى المحك » ليست المحكات كما شاعت فى مجال القياس النفسى ،
وانما هى المستويات ، وكان الاجدى لها أن تسمى « الاختبارات المرجعة
الى المستويات » حتى لا توقعنا فى هذا الخلط السيمانتى الذى لا مبرر له .
وعلى كل فقد أثرنا الإشارة إليها بالاصطلاح الذى شاع فى الوقت الحاضر
على أن يكون القارئ واعيا طول الوقت فى أن ما يقصد بالمحك فى هذا النوع
من الاختبارات هو المستوى .

أما المحكات حين تستخدم فى الحكم على الأداء الاختبارى وتفسيره ،
فإن ذلك توسيدها لمعناها فى بحوث الصدق وامتدادا له الى عالم الواقع ،
ولعل ما يتفق مع هذا المعنى ما اقترحه فؤاد أبو حطب من إشارة الى الصدق
(التقويم النفسى)

المرتبط بالمحككات بالصدق العملى التطبيعى (*) . ويتضح ذلك من عرض أشهر الطرق التى تستخدم فيها المحككات فى الحكم على الدرجة الخام فى الاختبارات النفسية وتفسيرها وهى ما يسمى طريقة جداول التوقع . Expectancy tables .

جداول التوقع : سبق أن قلنا أن كلا من الصدق التنبؤى والصدق التلازمى يعتمد على تقويم درجات الاختبار فى ضوء محك مستقل يهدف الاختبار الى التنبؤ به أو تشخيصه على التوالى . والعلاقة بين درجات المفحوصين فى الاختبار ومكانتهم فى المحك يمكن أن يحللها الباحثون بطرق مختلفة ، وأبسط صورة للتعبير عن هذه العلاقة هى ما يسمى جدول التوقع ومثل هذا الجدول يوضح توقع احتمال حصول الفرد على درجة معينة فى المحك (وليكن النجاح المدرسى أو المهنى) من حصوله على درجة معينة فى الاختبار .

وقد يأخذ جدول التوقع صورة رسم بيانى للمحك ذى البعدين (نجاح أو فشل مثلا) يوضح العلاقة بين النسب المئوية للأشخاص الذين استبعدوا من التدريب المبدئى على الطيران وبين الدرجة التساعية التى يحصل عليها الشخص فى بطارية انتقاء الطيارين التى أعدها السلاح الجوى الأمريكى . ويلاحظ من هذا الجدول أن ٧٧٪ من الأشخاص الذين حصلوا على التساعى الأول استبعدوا بينما نجد أن حوالى ٤٪ فقط من أولئك الذين حصلوا على التساعى التاسع فشلوا فى اتمام التدريب . كما يتضح لنا من الجدول أن النسب المئوية تتزايد تدريجيا فى التساعيات التى تقع بين الأول والأخير .

والنسب المئوية المعطاة فى مثل هذا الرسم تعطى تقديرا لاحتمال حصول الأفراد الذين تختبرهم على مكانة معينة فى المحك . فنحن نقتبأ مثلا من هذا الشكل أن حوالى ٤٠٪ من طلاب الطيران الذين يحصلون على الدرجة التساعية ٤ سوف يفشلون وأن حوالى ٦٠٪ منهم سوف ينهون برنامج

(*) استخدم فؤاد أبو حطب هذا المصطلح فى سلسلة محاضراته التى ألقاها عامى ١٩٧٩ - ١٩٨٠ ، ١٩٨٠ - ١٩٨١ بالمركز القومى للبحوث الاجتماعية والجنائية بالقاهرة عن الجوانب السيكومترية للاختبارات الاسقاطية .

التدريب على أعمال الطيران بنجاح . ويمكن تقرير توقعات مشابهة بنجاح أو فشل الأفراد الذين يقعون في أى من المعايير التساعية . وعلى ذلك فالفرد الذى يحصل على درجة تساعية مقدارها ٤ تكون فرصته $\frac{3}{4}$ أى $\frac{2}{3}$ فى انهاء برنامج التدريب على أعمال الطيران بنجاح .

وعندما يكون المحك متصلا وليس من النوع ذى البعدين فحسب فان جدول التوقع يمكن اعداده اعدادا مباشرا من الرسم البيانى للانتشار ، وفى اعداد رسم الانتشار يحدد لكل فرد درجته فى الاختبار وفى المحك وفيه يمثل الخط الأفقى مقاييس المحك (تقديرات التلاميذ فى مادة البلاغة) ويمثل الخط الرأسى الدرجات فى اختبار الجمل . وفى هذا الاختبار يطلب من المفحوص أن يحدد الأخطاء فى النحو أو علامات التوقيف أو استعمال الكلمات فى سلاسل من الجمل . وكل علامة فى الشكل تدل على كل من الدرجة فى الاختبار والتقدير فى المحك لكل طالب من مائة ، كما ان الرسم يتضمن العدد الكلى لتكرارات فى الخلايا المختلفة وكذلك المجموع الخام للتكرارات .

ولتحويل جدول الانتشار هذا الى جدول توقع تحول التكرار الموجود فى كل خلية الى نسبة مئوية من المجموع الخام المقابل له . فمثلا يوجد ١٦ شخصا حصلوا على الدرجات بين ٢٠ و ٢٩ فى الاختبار ، من بينهم ٦٪ (شخص واحد) حصل على التقدير (هـ) فى مقرر البلاغة و ١٩٪ (٣ أشخاص) حصلوا على التقدير (د) و ٥٦٪ (٩ حالات) حصلوا على التقدير (حـ) و ١٩٪ (٣ أشخاص) حصلوا على التقدير (ب) . وبنفس الطريقة يمكن ان تحول جميع تكرارات الخلايا فنحصل على جدول التوقع الآتى :

جدول (٧)

جدول توقع يوضح العلاقة بين الدرجات فى اختبار الجمل
وبين تقديرات مادة البلاغة

النسب المئوية للأفراد الذين حصلوا على كل تقدير					درجات الاختبار
ا	ب	ج	د	هـ	
١٠٠					٨٩ - ٨٠
٨٠	٢٠				٧٩ - ٧٠
٢٣	٦٣	١٤			٦٩ - ٦٠
٢٦	٣٥	٣٩			٥٩ - ٥٠
	٢٧	٥٩	١٤		٤٩ - ٤٠
	١٩	٥٦	١٩	٦	٣٩ - ٣٠
		٣٧	٥٠	١٣	٢٩ - ٢٠
			١٠٠		١٩ - ١٠
			١٠٠		٩ - صفر

ومن هذا الجدول يمكن التنبؤ بالتقدير الذى يحصل عليه الفرد فى المحك من الدرجة التى يحصل عليها فى الاختبار . فمثلا يمكن أن نتوقع أن الشخص الذى يحصل على درجة أعلى من ٨٠ أن يحصل على تقدير (ا) والذى يحصل على درجة أعلى من ٧٠ اما أن يحصل على تقدير (ا) أو (ب) ، كما نتوقع أن الشخص الذى يحصل على درجة أعلى من ٥٠ أن يحصل على تقدير (هـ) أو (د) . وهناك عديد من التوقعات التى يمكن الوصول اليها من هذا الجدول .

ولا بد أن نشير فى هذا الصدد أن مثل هذه التنبؤات تتعرض لأخطاء العينة وخاصة اذا كان عدد الأفراد الذين نختبرهم صغيرا . وحيث أن النسب المئوية الفردية تعتمد على عدد صغير نسبيا من الأفراد فى كل خلية فإن ذلك يعنى أن التغيرات الناتجة عن الصدفة عند الانتقال من عينة لأخرى قد تكون كبيرة .

المستويات :

النوع الثالث من الأسس التى تستخدم فى الحكم على الأداء الاختبارى

وتفسيره ما يسميه فؤاد أبو حطب (٩) المستويات standards وهي - كما ذكرنا انفا تتشابه مع المعايير في أنها أسس للحكم على الأداء في ضوء هذا الأداء ذاته وليس في ضوء أسس خارجية عنه ، إلا أنها تختلف عن المعايير في أنها قد تأخذ الصورة الكمية أو الكيفية ، وفي أنها تتحدد في ضوء ما يجب أن يكون عليه الأداء الاختباري وليس ما هو عليه بالفعل (كما هو الحال في المعايير) .

ولعل هذا النوع هو ما يقصده بناة ما يسمى في الوقت الحاضر الاختبارات المرجعة الى المحك ، « فالمحك » المستخدم في هذه الاختبارات ليس أساسا « خارجيا » وإنما هو أساس « داخلي » من بنية الاختبار ذاته ولعل أول مع أوقع حركة الاختبارات النفسية في هذا الخلط روبرت جليزر R. Glaser ، أول من استخدم هذا المصطلح عام ١٩٦٢ دون أن يفتبه الى استخدام مصطلح « محك » في القياس النفسي بمعنى ثابت ومحدد . ومما يؤسف له حقا أن سار على منواله عدد كبير من قادة القياس النفسي أنفسهم ، وحينما وجدوا أنفسهم في حيرة من أمرهم حاولوا البحث عن حلول توفيقية لعل أشهرها محاولة انستازي في طبعة ١٩٨٢ من كتابها الشهير (١١ : ٩٤ - ١٠١) .

ونعرض فيما يلي استخدام المستويات في أنواع مختلفة من الاختبارات النفسية :

١ - مستوى نطاق المحتوى : لعل أقوى ما يدعم وجهة نظرنا في استخدام مصطلح مستوى بدلا من مصطلح محك كأساس تفسيري للأداء في « الاختبارات المرجعة الى المحك » ما تتطلبه طبيعة هذه الاختبارات ومعظمها من نوع الاختبارات التحصيلية من الرجوع - في الحكم والتفسير على الأداء الاختباري - الى نطاق المحتوى content domain بدلا من الرجوع الى خصائص عينة تقنين معينة (كما هو الحال في المعيار) أو الى خصائص أداء مستقل عن الاختبار (كما هو الحال في المحك) ، ولهذا يشار اليها أحيانا بالاختبارات المرجعة الى المحتوى أو النطاق أو الأهداف ، وهذا ما تفضله انستازي (١١ : ٩٤) ، لأن مصطلح « الاختبارات المرجعة

الى المحك ، - على الرغم من شيوعه - ليس افضل ما يشير الى هذا النوع من الاختبارات .

ويتم الحكم على الأداء الاختباري وتفسيره هنا في ضوء معنى المحتوى حيث يتم التركيز على ما يؤديه المفحوص بالفعل مقارنة بما يجب ان يؤديه ، وليس مقارنة بأداء الآخرين . ويتطلب ذلك ان يتحدد مقدما نطاق المعرفة او المهارة التي يقيسها الاختبار . ومن المطالب الجوهرية في هذه الحالة ان يكون ، نطاق المحتوى « هذا الذي يتم اختياره - او عينة منه - في بناء الاختبار له أهمية ، ثم يقسم النطاق المختار الى وحدات صغيرة في صورة عبارات أدائية . وهذا ما يسمى في السياق التربوي صياغة الأهداف التعليمية صياغة سلوكية اجرائية (راجع ، ٧) .

والهدف الجوهرى للمقارنة في هذه الحالة هو الحكم على مدى الاتقان mastery وهو جوهر الحكم في هذه الحالة ، وهو من نوع الكل او لا شيء ، ويشير الى مدى وصول المفحوص الى المستوى المحدد مقدما للاتقان ، وقد تستخدم ثلاثة مستويات للحكم على مستوى الاتقان ، المستوى المتوسط او موضع الشك او الذى يحتاج لمراجعة ، ومستوى عدم الاتقان او الفشل ويحتاج هذا بالطبع الى تحديد نقطة القطع cutoff point التى عندها يمكن ان تصدر مثل هذا الحكم . وتوجد طرق مختلفة لتحديد نقاط القطع يدرسها بالتفصيل محمود محمد ابراهيم عبد الله فى رسالته للماجستير بكلية التربية جامعة الأزهر (١١) .

ويحتاج الوصول الى الحكم فى ضوء المستوى استخدام اجراءات القرار التتابعى sequential decision وفيها لا يعرض على المفحوص عدد ثابت محدد مقدما من الأسئلة - كما هو الحال فى الاختبارات المعتادة ، وإنما يستمر المفحوص فى اخذ الاختبار حتى يصل الفاحص الى نقطة القطع التى عندها يصل الفاحص الى اتخاذ قرار باتقان المفحوص للمحتوى او عدم اتقانه له . وعند هذه النقطة يتوقف الاختبار ويوجه المفحوص اما الى المستوى التعليمى التالى او يعود الى تعلم المحتوى نفسه من جديد . وقد اصبحت هذه الاجراءات التتابعية أكثر يسرا مع استخدام الحاسبات

الالكترونية. التى تختصر وقت الاختبار عن ناحية وتوصل الى تقدير للاتقان
أكثر دقة .

٢ - المعايير الارتقائية الرتبة : تعود هذه المعايير بأصولها الى نتائج
بحوث علم نفس النمو والتي أدت الى وصف السلوك الانسانى فى مراحل
المتابعة ومن أشهر الاعمال المبكرة فى هذا المجال جهود جيزل وزملائه فى
جامعة ييل بالولايات المتحدة والتي نتجت عنها جداول جيزل الارتقائية
Gesell Developmental Schedules وتوضح هذه الجداول المستوى
النمائى التقريبى التى يصل اليه الطفل مقدرا بالشهور فى أربعة مجالات
رئيسية هى النشاط الحركى والتوافقى واللغوى والشخصى الاجتماعى .
ويتحدد هذا المستوى بمقارنة سلوك الطفل وبالسلوك المعتاد للأطفال من
سن ٤ أسابيع وحتى ٣٦ شهرا ونشأت عن ذلك مقاييس ذات طبيعة رتيبة
متضمنة ثلاثة افتراضات أساسية هى :

١ - يتسم السلوك المبكر فى نموه بالتتابع المنتظم .

٢ - تتتابع المراحل النمائية الارتقائية لسلوك الانسان بنظام ثابت .

٣ - تعتمد كل مرحلة لاحقة على اتقان الأنماط السلوكية اللازمة
للمراحل السابقة .

ومنذ مطلع الستينات شهد الميدان اهتماما كبيرا بالنظريات الارتقائية
لعالم النفس السويسرى العظيم جان بياجيه ، والتي توصل اليها من خلال
برنامج طموح للبحث فى النشاط المعرفى عند الاطفال والمراهقين بداء منذ
منتصف العشرينات . وقد وظفت المهام التى استخدمها فى معظم الدراسات
اللاحقة فى علم نفس النمو ، كما بنيت اختبارات عقلية مقننة على أساسها
وتعد هذه الاختبارات من النوع الرتبى أيضا وتصدق عليها الافتراضات
السابقة .

وللتمييز بين هذا النوع من المعايير ومعيار العمر العقلى نشير الى أنه
فى المعايير الرتبية تصمم المهام (التى قد تتحول الى أسئلة أو مفردات فى

(الاختبار) بحيث تكشف عن الجوانب الهامة والمباعدة لكل مرحلة ارتقائية،
ثم تجمع بعد ذلك البيانات الامبريقية حول الاعمار التى يتم التوصل فيها
الى كل مرحلة من هذه المراحل . اما فى معيار العمر العقلى فالعكس هو
الصحيح ، ففيه تختار المهام (الأسئلة) على اساس تمييزها بين الأعمار
المتتابة . ويوجد وجه آخر للاختلاف بين نوعى المعايير هو أن معيار العمر
العقلى كمى فى جوهره ويتحدد فى ضوء درجة يحصل عليها المفحوص فى
الاختبار . اما فى المعيار الارتقائى الرتبى فان الأهم هو الوصف الكيفى
لسلوك الطفل ولذلك فهو يعتمد على المستوى وليس المعيار .

مراجع الباب الثانى

- ١ - أحمد الرفاعى محمد غنيم : تطبيقات على ثبات الاختبارات ، القاهرة ، مكتبة نهضة الشرق ، ١٩٨٥ .
- ٢ - رمزية الغريب : التقويم والقياس النفسى والتربوى ، القاهرة : مكتبة الأنجلو المصرية ١٩٧٠ .
- ٣ - سيد أحمد عثمان : لمحة الى ما وراء المنهج : مقدمة الطبعة الثانية لكتاب (مناهج البحث فى التربية وعلم النفس) تأليف د.ب. فان دالين (ترجمة محمد نبيل : فل وآخرين) مكتبة الأنجلو المصرية ، ١٩٧٩ .
- ٤ - _____ : الموضوعية والذاتية ، الكتاب السنوى فى التربية وعلم النفس (تحرير سعيد اسماعيل على) ، المجلد الخامس ، ١٩٧٨ .
- ٥ - صلاح قنصوة : الموضوعية فى العلوم الانسانية ، دار الثقافة للطباعة والنشر ، ١٩٨٠ .
- ٦ - صلاح مخيمر : عن الذاتية والموضوعية فى علم النفس ، مكتبة سبيد رافت ، (ب.ت) .
- ٧ - فؤاد أبو حطب ، آمال صادق علم النفس التربوى (ط ٣) ، مكتبة الأنجلو المصرية ، ١٩٨٤ .
- ٨ - فؤاد أبو حطب قضية سيرل بيرت ، محاضرة عامة القيت بالجمعية المصرية لدراسات النفسية ٢ سبتمبر ١٩٧٨ .
- ٩ - فؤاد أبو حطب ، بديوى علام : الاحصاء النفسى والتربوى ، مكتبة الأنجلو المصرية (تحت الطبع) .
- ١٠ - محمد خليفة بركات : مشروع مبدئى لثقاق اخلاقيات المهنة للمشتغلين

بعلم النفس - المؤتمر السنوى الثانى لعلم النفس ، الجمعية المصرية
للدراسات النفسية ، أبريل ١٩٨٦ .

١١ - محمود محمد ابراهيم عبد الله : دراسة مقارنة لبعض طرق حساب
معامل الاعتمادية لنقاط القطع فى الاختبارات مرجعية المحك . رسالة
ماجستير - كلية التربية جامعة الأزهر ١٩٨٦ .

11. Anastasi, A. Psychological Testing. (3rd. ed.) New York : Macmillan, 1968, 1976, 1982.
12. A.P.A. Ethical Standards for psychoogists. Washington, D.C., 1953.
13. A.P.A. Technical Recommendations for psychological Tests and Diagnostic Techniques, Washington D.C. 1954.
14. A.P.A. Standards for educational and psychological tests and manuals. Washington, D.C. 1966, 1974.
15. A.P.A. Ethical Standards for Psychologists Amer. Psychologist, 1968, 23, 357-361.
16. ——— Casebook on ethical statements of psychogists. Washington, D.C., 1974.
17. ——— Ethical principles in the conduct of research with human participants, Washington, D.C., 1973.
18. Benton, A. L. Influence of incentives upon intelligence test scores of school children, J. Genet. Psychol., 1936, 49, 494, 496.
19. Borden, E.S. Psychological Counseling, New York, Appleton, 1965.
20. Cronback, J.L. Essentials of Psychological Testing (3rd ed.) New York : Harper, 1960, 1970, 1984.
21. Cronbach, J.L., Meehl, P.E. Construct validity in psychological tests. Psychol. Bull., 1955, 52, 281-302.
22. Flanagan, J.C. The development of an index of examine motivation Educ. Psychol. Measurement, 1955, 15, 144-161.
23. Ghiselli, G.E., Campbell, J.P. and Zadeck, S. Measurement theory for the behavioural sciences, San Francisco : W.H. Freeman, 1981.
24. Guilford, J.P. Psychometric methods. New York : McGraw-Hill, 1954.
25. Hurlock, E.B. An evaluation of certain incentives used in school Work. J. Educ. Psychol., 1925, 16, 154-159.

26. Lindquist, E.F. (ed.) Educational measurement Washington, D.C., Amer. Council on Educ., 1951.
27. Nunnally, J.C. Psychometric theory. New York : McGraw-Hill, 1978.
28. Ronan, W.W. & Prien, E.P. Perspectives on the measurement of human performance. New York : Appleton-Century-Crofts, 1971.
29. Sarason, S.B. et al., A test anxiety Scale for children. *Child Develop.*, 1959, 29, 105-115.
30. Stanley, J.C. & Hopkins, K.D. Educational and Psychological measurement evaluation. N.J. : Prentice-Hall, 1972.
31. Thorndike, R.L. (ed.) Educational Measurement (2nd ed.), Washington : D.C. Amer. Council on Educ., 1971.
32. Thorndike, R.L. & Hagen, E. Measurement and evaluation in psychology and education (4th ed.) New York : Wiley, 1977.

الباب الثالث
اختبارات الذكاء العام

تقديم للباب الرابع

توجد فئة من الاختبارات النفسية تسمى اختبارات الذكاء العام .
وتتميز هذه الاختبارات بأنها تستخدم في مواقف متنوعة ويتحدد صدقها في ضوء مدركات أكثر شمولاً . كما تتميز هذه الاختبارات بأنها تعطي درجة واحدة مثل نسبة الذكاء لتدل على المستوى العقلي العام للمفحوص . ويرى ترمان وميرل (٥٦) أن الوصول إلى هذا التقدير الشمولي للمستوى العقلي يتطلب خبرة من الباحث في «تصويب الرماح إلى الأهداف الحاسمة» . ولتحقيق ذلك عليه أن يعرض على المفحوص عدداً كبيراً ومتنوعاً من الأعمال على أساس أنها تمثل عينة ملائمة من أساليب الأداء العقلي . ومع ذلك فإن البحوث الحديثة تؤكد أن هذه الاختبارات تتشبع بعمليات عقلية معينة مثل القدرة اللغوية ، وتتجاهل تماماً بعض العمليات العقلية الأخرى ، كقدرة الموسيقى .

وكثيراً ما تسمى اختبارات الذكاء العام باختبارات الاستعداد المدرسي Scholastic aptitude لأن صدقها يتحدد في العادة في ضوء محكات التحصيل الأكاديمي . كما أن هذه الاختبارات تستخدم في أغلب الأحوال في أغراض التصفية المبدئية العامة (سواء من الناحية التربوية أو المهنية أو العسكرية) ، والمفروض أن يتبعها تطبيق اختبارات الاستعدادات الأكثر تخصصاً . ويشيع هذا على وجه الخصوص في اختبار المراهقين والراشدين الأسوياء عند الإرشاد أو الانتقاء أو التصنيف أو التوجيه أو غير ذلك من الأغراض التي أشرنا إليها في الفصل الأول من هذا الكتاب . كما يوجد استخدام آخر لاختبارات الذكاء ، في الأعراض الكلينيكية خاصة ، وهو تحديد الضعف العقلي وتشخيصه . وفي مثل هذه الأحوال فإن استخدام الاختبارات الفردية أكثر شيوعاً من غيره ، إلا أننا كثيراً ما نستخدم الاختبارات الجماعية للذكاء في أغراض البحث والممارسة السيكولوجيتين .

وسوف نخصص في هذا الباب لاختبارات الذكاء العام فيتناول الفصل السابع اختبارات الذكاء الفردية أما الفصل الثامن فسوف يتبادل الاختبارات الجماعية للذكاء .

الفصل السابع

اختبارات الذكاء الفردية

يتناول هذا الفصل بالعرض اختبارات الذكاء الفردية ، وعلى وجه الخصوص أشهر اختبارين فرديين للذكاء وهما : مقياس ستانفورد - بينيه ، ومقياس وكسلر ، ويعد مقياس بينيه على وجه الخصوص الوالد الشرعى لحركة قياس العقل الانسانى - وخاصة عند الأطفال ، ولذلك سيحظى فى هذا الفصل باهتمام خاص ، ثم نعرض لمقاييس وكسلر ، وبعض الاختبارات والمقاييس الأخرى الأقل شهرة والتي لا تقل أهمية .

مقياس ستانفورد - بينيه

الفرد بينيه :

الفرد بينيه (١٨٥٧ - ١٩١١) هو العالم الفذ الأصل ، بدأ تدريبه فى ميدان الطب ثم سرعان ما أصبح أشهر علماء النفس فى فرنسا فى عصره ؛ وله فضله العظيم على علم النفس فى بلده . فقد أنشأ مع زميله بيرونى Beauns أول معمل لعلم النفس عام ١٨٨٩ فى جامعة السوربون ، وأول مجلة متخصصة باللغة الفرنسية ، حولىة علم النفس L'Année Psychologique عام ١٨٩٥ .

ولقد بدأ بينيه عالما تجريبيا ، فقبل ان يعد اختباراتة العقلية المشهورة اهتم بدراسة العمليات العقلية المعرفية ، واطلق على منهجه اسم «سيكولوجية الفرد» . وفى رأيه ان أهم ما نجده من فوارق بين الناس انما يكون هو فى الأنشطة العقلية العليا كالاستدلال والحكم ، وهى جميعا من العمليات التى يصعب قياسها بالوحدات الفيزيائية (كالأحساس) ، كما لا تتوافر فيها الوحدات الأولية للنشاط الذهنى ، والتي كان يهتم بها أغلب علماء النفس من انصار بنائيه فوندت (ومنهم ، تيتشر) (١٤) .

وقد تأكدت صحة وجهة نظر بينيه من نتائج بعض الدراسات التى

(التقويم النفسى)

أجريت على هـ صدق اختبارات جيمس ماكس كاتل (الذى اقتفى فيها اثر السير فرنسيس جالتون فى إنجلترا) لمقاييس عقلية ، وبخاصة بحوث كلارك وزلر وشتيلا شارب (٢٤) ، ولذلك بدأ بينيه يطبق فى بحوثه عن الاختبارات العقلية ما توصل اليه من نتائج البحث التجريبي فى عملية التفكير ، فانتقد اختبارات كاتل وجالتون لأنها تقتصر على النواحي الحسية البسيطة ، وفضل الاختبارات الأكثر تعقدا وتركيبا ، واقترح عددا من الاختبارات لقياس « الملكات » العقلية (٥٠) .

وقضى بينيه الفترة بين عامى ١٨٩٦ و ١٩٠٥ فى اجراء سلسلة طويلة من البحوث على الأطفال الأسوياء وغيرهم مطبقا أنواعا عديدة من الاختبارات لقياس مختلف « الملكات » العقلية ، بالإضافة الى عدد من التجارب أجراها حول المسائل الخلافية فى سيكولوجية التفكير فى ذلك الوقت مثل التجريد ، والتفكير بدون صور وزمن الرجوع والتمييز الحسى . وكان الهدف الذى يسعى اليه من ذلك هو تحديد العلاقة بين هذه الاختبارات والعمر الزمنى والتحصيل المدرسى وتقديرات المعلمين للقدرة العقلية ، وكانت بحوثه جيدة التصميم وأدت به الى التحول من الاهتمام بالاختبارات التى تقيس « الملكات » التقليدية الى تلك التى تقيس المستوى العقلى بالفعل .

وهكذا كانت نظرة بينيه الى الذكاء أكثر شمولاً مما ظهر فى مقياسه ذاته بعد ذلك ؛ حين لجأ الى استخدام درجة كلية واحدة هى العمر العقلى ، فى قياس الذكاء . ويمكن للقارئ المهتم باسهامات بينيه فى النظرية الأساسية للذكاء أن يرجع الى (٢٤) .

نشأة مقياس بينيه :

فى عام ١٩٠٤ طلبت وزارة المعارف الفرنسية من الفرد بينيه والطبيب الفرنسى ثيودور سيمون دراسة الطرق التى يمكن أن تستخدم فى تصنيف الأطفال الذين يتميزون ببطء التعلم فى المدارس الفرنسية ، فعكفا على وضع مقياسهما المشهور للذكاء بغرض التمييز بين الأطفال الأسوياء وضعاف العقل .

ولم يكن الهدف دراسة وتحليل استعدادات المتخلفين عقليا ، وانما كان الهدف عمليا بحثا وهو تقويم وقياس ذكاء هؤلاء الاطفال بصفة عامة وتحديد مستواهم العقلي ومقارنتهم بالاطفال الأسوياء من نفس العمر والمستوى ١٠ أى الهدف كما يقول بينيه وسيمون (٥٠) هو بناء مقياس « مترى » Metrical للذكاء يتكون من عدد من « الاختبارات » المتدرجة فى الصعوبة ، مرتبة حسب الأعمار حتى يستطيع الفاحص أن يحدد مدى تقدم الطفل عقليا أو تخلفه . وقد سعى بينيه الى أن تكون هذه « الاختبارات » متنوعة ، مع التركيز على فئات معينة من العمليات العقلية اهتم بها دون غيرها عند اعداد المقياس وهى : الحكم والفهم العام والمبادأة والقدرة على التكيف .

ويتكون مقياس بينيه الأصيل الذى نشر عام ١٩٠٥ باسم مقياس بينيه - سيمون للذكاء Binet-Simon Scale من ٢٠ سؤالا أطلق على كل منها اسم « اختبار » وشاع هذا الاسم فى جميع الطبقات التالية لهذا المقياس فى جميع اللغات (*) ، وفيما يلى قائمة بهذه الاسئلة ، نذكرها كاملة لنبين الى أى حد اختلفت عن اهتمامات علماء القياس العقلي قبل بينيه ، وكيف أنها هيأت الموقف لظهور عدد كبير من الاختبارات العقلية تنحدر فى معظمها « من صلب » مقياس بينيه على حد تعبير أنستازى (٤٩) بالاضافة الى شيرورها فى معظم تعديلات مقياس بينيه ذاته .

- (١) التأزر البصرى (تتبع شمعة مضيئة بالرأس والعينين) .
- (٢) الفهم عن طريق ولمس (الامساك بمكعب بعد لمسه) .
- (٣) الفهم عن طريق البصر أو الرؤية .
- (٤) التعرف على الطعام (التمييز بين قطعة الخشب وقطعة الشيكولاته) .
- (٥) البحث عن الطعام (الاستجابة لقطعة من الشيكولاته ملفوفة فى الورق) .
- (٦) تتبع التعليمات البسيطة أو الأيماءات المتكررة .

(*) أثرفنا أن نشير الى « الاختبارات » التى يتألف منها مقياس ستانفورد - بينيه بمصطلح « الاسئلة » فى متن هذا الكتاب ، حتى لا تختلط بمعنى كلمة « الاختبار » كما تشيع فى التراث السيكولوجى الحديث .

(٧) الإشارة الى الأشياء وخاصة أعضاء الجسم (الرأس ، الأنف ، الخ) .

(٨) التعرف على الأشياء من صورها .

(١٠) التمييز الحسى بين طول خطين .

(١١) تكرار ثلاثة أرقام .

(١٢) التمييز بين وزن شيئين .

(١٣) القابلية للإيحاء المقارنة بين خطين متساويين - أو البحث عن

أشياء لا وجود لها فى صورة) .

(١٤) تعريف الكلمات البسيطة (التى تدل على أشياء مألوفة) .

(١٥) تكرار جملة تتكون من ١٥ كلمة .

(١٦) بيان أوجه الاختلاف بين شيئين .

(١٧) الذاكرة البصرية .

(١٨) رسم الأشكال من الذاكرة .

(١٩) مدى ذاكرة الأرقام .

(٢٠) بيان أوجه التشابه بين الأشياء .

(٢١) التمييز بين الخطوط بسرعة .

(٢٢) ترتيب ٥ أوزان .

(٢٣) تحديد الوزن الناقص (من بين الأوزان الخمسة فى الاختبار

السابق) .

(٢٤) إعطاء كلمات ذات قافية واحدة .

(٢٥) تكملة الجمل .

(٢٦) بناء جملة تحتوى على ٣ كلمات معينة .

(٢٧) الإجابة على الأسئلة (مثل : ماذا تفعل وانت نعلان ؟) .

(٢٨) تحديد الوقت بعد أن تحل عقارب الساعة محل بعضها .

(٢٩) ثنى وتقطيع قطعة من الورق .

(٣٠) التمييز بين الكلمات المجردة (مثل ، حزن وسأم) .

وإذا فحصنا هذه القائمة فحصا دقيقا نجد أن الأسئلة الثلاثة الأولى

هى فى الواقع أسئلة للنمى الحركى ، أما الأسئلة الأخرى فيمكن أن تعد من

النوع العقلى ، ومن بينها ١٨ سؤالا (أى حوالى الثلثين) من النوع الذى

يقيس ما يسميه جيلفورد التفكير المعرفة و ٥ أسئلة تقيس ما يسميه الذاكرة

بأنماطها المختلفة (الأسئلة ١١ ، ١٥ ، ١٧ ، ١٨ ، ١٩) ولا يوجد الا ثلاثة أسئلة فقط تقيس ما أطلق عليه فيما بعد قدرات التفكير الابتكارى (او الانتاج التباعدى عند جيلفرورد) .

وكانت الطريقة التى استخدمها بينيه فى تقنين مقياسه غير ملائمة ، فقد طبقه على ٥٠ طفلاً (١٠ من كل فئة من فئات الاعمار ٢ ، ٥ ، ٧ ، ٩ ، ١١) اختارهم معلموهم على أساس أنهم متوسطو القدرة العقلية ، بالإضافة الى عدد من الاطفال لم يذكره بينيه كانوا من ضعاف العقول . وكانت طريقة التصحيح تعتمد على أساس تجريبي مبتدى يسمح بتصنيف الاطفال تبعاً لأعلى مستوى يصل اليه الطفل ؛ فالمعتوهون لا يمكنهم تجاوز السؤال السادس ، ولا يتجاوز البلاء السؤال الخامس عشر ، أما حدود الماقونين فلم تكن معروفة ، رغم أنهم يعجزون فى العادة عن الاجابة على الأسئلة المجردة التى يستطيع معظم الاطفال العاديين من سن ١١ الاجابة عليها .

وبالنطبع ارتبط عمر المفحوصين بنجاحهم فى الاختبار ، ولم تكن لدى بينيه حتى ذلك الوقت (عام ١٩٠٥) طريقة تسمح بالمقارنة بين الأعمار . ولر أنه كان يلتمس طريقه فيما يبدو نحو مفهزم العمر العقلى . يقول بينيه وسيمون عام ١٩٠٥ (٥٠ : ٢٢١) .

ومن المحتمل أن ننجح فى يوم ما فى ايجاد علامات على التخلف السيكولوجى مستقلة تماماً عن العمر . وقد يكون من المفيد جداً التعرف على هذه العلامات ، الا أننا فى الوقت الحاضر نجد ان ما يجذب انتباهنا اثر من غيره هو التشابه بين الأسوياء من صغار السن وغير الأسوياء من كبار السن . وهذا التشابه من الكثرة والغرابية بحيث أننا لو حصلنا على استجابات طفل لا نعرف عمره لا نستطيع الحكم عليه بالسواء أو عدم السواء .

وكانت لطبعة عام ١٩٠٥ من مقياس بينيه عدة سمات تعد نقطة تحول فى تاريخ القياس العقلى ففيه نجد لأول مرة ان الأسئلة التى تقيس مكونات عديدة ترتبط فيما بينها لتحدد مستوى من الأداء بدلا من القياس المنفصل لأبعاد التحليل السيكولوجى التقليدى الذى كان شائعاً فى عصر بينيه مثل

الملكات أو عتبات الاحساس أو غيرها وبالإضافة الى ذلك فإن وضع سؤال فى المقياس أو حذفه أو تعديله كان من المسائل التجريبية ، ويدل على خبرة بينيه الطويلة بالتجريب العملى . فقد كان يهتم بمدى العلاقة بين السؤال وأحكام المحك (أى بنوع من الصدق) . ثم ان هذه الاختبارات لم تكن محددات معملية دقيقة وإنما كانت أدوات عملية بسيطة وهامة لدراسة الأطفال فى مواقف واقعية (مثل حجرة الدراسة) .

وتد أدى ظهور مقياس ١٩٠٥ الى جذب انتباه عدد كبير من المهتمين بمشكلات الضعف العقلى والتصنيف التربوى ، فقام العالم الأمريكى جودارد Goddaard بترجمته الى الانجليزية ، كما جـسـريـه ديكرولى Decroly وديجان Degand فى بلجيكا . وكان الكثيرون - ومنهم بيرت فى انجلترا وترمان فى أمريكا - على اتصال مباشر ببينيه كما اقترحوا فى نفس الوقت اختباراتهم الخاصة .

تعديل عام ١٩٠٨ :

فى عام ١٩٠٨ ظهر مقياس جديد لبينيه وسيمون يعتمد اعتمادا كبيرا على المقياس السابق ومع تعديلات جوهرية ، فقد أصبح اسم المقياس الجديد « نمو الذكاء عند الأطفال ، The development of intelligence in children » وتحول تركيز هذا المقياس من غير الأسوياء الى الأسوياء ، واستبعدت اختبارات المعتوهين . واحتفظ المقياس بخاصية استخدام الأسئلة القصيرة ، العديدة المتنوعة من حيث المحتوى ، والمرتبة ترتيبيا تجريبيا حسب الصعوبة . ولأول مرة تصنف الأسئلة حسب مستويات الأعمار ابتداء من سن ٣ سنوات حتى سن ١٣ سنة تبعا للعمر الذى يستطيع الأطفال الأسوياء الوصول الى أسئلته بنجاح . وكان تحديد الأسئلة للمستويات العمرية المختلفة تحديدا جزافيا ، ليس بسبب قلة عدد أفراد عينة التقنين فحسب وإنما أيضا بسبب الصعوبات العملية فى اعداد الأسئلة التى تقع فى مستويات الصعوبة المطلوبة ، وبوجه عام اعتبر المستوى العمرى للاختبار هو ما يستطيع ما بين ٥٥٪ - ٧٥٪ من أطفال عينة التقنين بالاجابة عليه اجابة صحيحة . لأن هذا المدى يتضمن نسبة الـ ٥٠٪ التى تدل دلالة عامة على الأداء المتوسط وكانت عينة التقنين ٢٠٢ أطفال تتراوح أعمارهم بين ٣ سنوات و ١٣ سنة .

وقد أدى ترتيب الأسئلة النمائي حسب مستويات العمر بالفاحصين الى التعبير عن المستوى الارتقائي للطفل بالعمر الذى يحصل فيه الطفل المتوسط على درجة مكافئة . ورغم أنه لم تتحدد مستويات النجاح والفشل على أساس أسئلة المقياس ، فإن العمر العقلى للطفل كان يتحدد على أساس تساوى العمر الذى يستطيع أن يجتازه ويجتاز الأعمار السابقة عليه (العمر القاعدى أو الأساسى) ثم يضاف اليه عام اضافى من العمر العقلى لكل ٥ أسئلة من المستويات الأكثر صعوبة .

هذه هى الخصائص الأساسية لمعيار العمر العقلى كما اقترحه تعديل ١٩٠٨ ، وهو المعيار الذى اقتبسته مقاييس الذكاء اللاحقة . ورغم أن هذا المعيار يتضمن مشكلات احصائية اشرنا اليها فى الفصل السادس ، فإن له فائدته فى أنه يوفر للباحث طريقة سهلة الفهم فى وصف الفروق بين الأطفال فى المستوى العام للقدرة العقلية . ومع ذلك بقيت مشكلتان هامتان لم يحلها مقياس ١٩٠٨ : أولاً أن عدد الأسئلة فى كل مستوى عمرى يختلف ، ويمتد بين ٣ ، ٨ أسئلة ، وثانيهما أن درجة العمر العقلى لا تفهم كمعيار للتفروق أو التخلف الا اذا ارتبطت بالعمر الزمنى للطفل .

تعديل عام ١٩١١ :

فى عام ١٩١١ قام بينيه - قبيل وفاته المبكرة - بتعديل مقياسه ونشره باسمه منفرداً ، فأعاد ترتيب كثير من الأسئلة ، ووجد عددها فى كل مستوى عمرى فجعله خمسة أسئلة ثم أضاف عدداً آخر من الأسئلة (٥ لمستوى ١٥ سنة ، و ٥ لمستوى الراشد) وهكذا أمكن - بعد تساوى عدد الاختبارات فى كل مستوى - اعطاء كسور من العمر العقلى مقدارها سنة لكل اختبار يستطيع الفحوص الاجابة عليه بعد عمره القاعدى أو الأساسى . ومع ذلك فإن مفهوم العمر العقلى ظل على حاله ، أى لا يمكن استخدامه كمؤشر على المستوى العقلى مستقلاً عن العمر الزمنى ، وقد أدى ذلك بالعالم الالماني شترن

Stern عام ١٩١٢ الى اقتراح معيار النسبة العقلية أى

العمر الزمنى

وهو المعيار الذى استخدمه - بعد التعديل - ترمان كما سنوضح بعد قليل .
ويوضح الجدول رقم (٨) الأسئلة التى يتكون مقياس عام ١٩١١ ، ومن هذا
الجدول يتضح أن المقياس فى صورته الأخيرة على يد بينيه يتكون من ٥٤
سؤالاً ، وفيه حذفت أسئلة المستويات الخاصة بالأعمار ١١ ، ١٢ ، ٢٤ سنة
كما أن أسئلة سن ٤ كانت ٤ أسئلة فقط (٥٣) .

جدول رقم (٨) يبين الاختبارات التي يشكون منها مقياس يينية (تعديل عام ١٩١١)

أمثلة تعديل عام ١٩١١					المر
صورة (سهل) ٠	٦ مقاطع	عددان ٠	الإشارة	اسم العائلة	٣
بطاقة مقسمة	خطان	٣ أعداد	التسمية	الجنس	٤
الصباح والمساء	١٠ مقاطع	وزنان	المربع	٤ قطع نقود	٥
صورة (صعب)	تعريف الكلمات (سهل)	مقارنة الوجوه	المعين	١٣ قطعه نقود	٦
المد من ٢٠ إلى ١	الترتيب الثلاثي	الجين والشمال	٤ ألوان	٣ قطع نقود، ٣ أنصاف هذه القطع	٧
الأمثلة السهلة	الفروق (المحسوسة)	٥ أعداد	التاريخ	السمات للنافضة	٨
الأمثلة الصعبة	تعريف الكلمات (محسوس)	الثور	التفسير	٩ قطع نقود	٩
الجل المختلطة	السخافات	٥ أوزان	الرسم من الذاكرة	بناء الجمل (سهل)	١٠
الصورة	تعريف الكلمات (مجرد)	٦٠ كلمة في ٣ دقائق	الإيجاء	بناء الجمل (صعب)	١٢
الفلت المكسوس	٢٦ مقطاً	٧ أعداد	٣ إيفاعات	للمشكلات والنائل	١٥
	الدروق المجرودة	الفرق بين حرفين	إعادة التعبير	الورق المطوى	الرشد

مقياس ستانفورد - بينيه لترمان :

لم تكن محاولة جو دارد ترجمة ونقل مقياس بينيه الى اللغة الانجليزية هي المحاولة الوحيدة ، ففي عام ١٩١٠ نشر هوي Huey ووبل Whipple ترجمة له ، كما نشر فالين Wallin وكولمان Kuhlmann ترجمة اخرى في عام ١٩١١ . وفي عام ١٩١٢ بدا بيرت Burt نقل المقياس الى الانجليزية وتقنيته على عينة كبيرة من الاطفال الانجليز عندما عين اخصائيا نفسيا لمقاطعة لندن ، وظل يتابع مشروعه هذا طوال تاريخه العلمي كأستاذ لعلم النفس بجامعة لندن . بالاضافة الى نقل المقياس في اقطار عديدة بلغاتها مثل ايطاليا وألمانيا وروسيا والسويد وتركيا واليابان . وفي عام ١٩٢٠ ترجم احمد فكرى (٥) مقياس بينيه الاصلى الى اللغة العربية . وقدم له تعليمات مقتضبة ، وكان يعتقد خطأ أن المدرس - دون تدريب سابق - يمكنه تطبيق المقياس والحكم على ذكاء الطفل . ومع ذلك فان أكثر هذه المحاولات شهرة في تاريخ المقياس النفسى محاولة لمويس ترمان L.M. Terman (١٨٧٧ - ١٩٥٦) أستاذ علم النفس بجامعة ستانفورد . والتي عرف بها المقياس باسم مقياس ستانفورد - بينيه للذكاء ، وقد شاعت هذه التسمية أكثر من الاسم الاصلى للمقياس .

وقد ظهرت الطبعة الاولى لمقياس ستانفورد - بينيه عام ١٩١٦ . وادخل ترمان على المقياس الاصلى عددا من التعديلات والاضافات الجوهرية بحيث يمكن القول أنه اختبار جديد . فقد كان حوالى ثلث الأسئلة جديدا ، وعدل ترمان عددا من الأسئلة القديمة واعيد توزيعها على المستويات العمرية المختلفة أو حذفت . واعيد تقنين المقياس كله على عينة من الأمريكيين بلغ عددها ١٠٠٠ طفل و ٤٠٠ راشد . واعدت تعليمات مفصلة لتطبيق كل سؤال وتصحيحه ، واستخدام معيار نسبة الذكاء (تعديلا لمعيار النسبة العقلية لشترن وكولمان) في صورتها التي شاعت في المقياس النفسى وهي :

العمر العقلى

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠$$

العمر الزمنى

وقد كتبت مجلة المقتطف المصرية مقالا عن هذا المقياس عام ١٩٢٧ (٤٦) كما كتب عنه أمير بقطر عام ١٩٢٨ (١٠) . ونقله إلى اللغة العربية لأول مرة طبيب غير متخصص في علم النفس هو حسن عمر (١٥) . ثم ترجمه على نحو أدق اسماعيل القباني (٦) عام ١٩٢٨ .

وفي عام ١٩٣٧ ظهرت في الولايات المتحدة الطبعة الثانية المعدلة للمقياس ، وتتكون من صريحتين متكافئتين ، الصورة ل ، والصورة م . وفي هذه الطبعة زاد عدد أسئلة المقياس وأعيد تقنيته تماما على عينة جديدة جيدة السحب من المجتمع الأمريكي بلغ عددها أكثر من ٣٠٠٠ مفحوص تتراوح أعمارهم بين سن ١٥ سنة إلى سن ١٨ . واشتملت كل عينة عمرية على عدد متساو من الذكور والإناث ، وتم اختبار معظم المفحوصين من سن ٦ سنوات وما بعدها من تلاميذ المدارس ، أما عينات أطفال ما قبل المدرسة (عينات الأعمار ١ ، ٢ ، ٣ ، ٤ ، ٥ ، ٦ ، ٧ ، ٨ ، ٩ ، ١٠ سنة) فتم الحصول عليها بطرق مختلفة ، أيسرها تطبيق المقياس على أخوة أطفال وتلاميذ المدارس ، وقد روعى في اشتقاق العينات التوزيع الجغرافي والمستويات الاقتصادية والاجتماعية .

وقد ظهرت الطبعة التجريبية للصورة ل باللغة العربية ١٩٥٦ وقام بإعدادها محمد عبد السلام أحمد ولويس كامل مليك (٤٠) .

وفي عام ١٩٦٠ ظهرت طبعة ثالثة لمقياس ستانفورد - بينيه في صورة واحدة فقط تجمع أفضل الأسئلة من صورتى طبعة عام ١٩٣٧ . وبدون تقديم أى محتوى جديد أو صياغة أسئلة جديدة كان من الممكن حذف الأسئلة التى لا تساير العصر . وأعيد توزيع الأسئلة التى تغيرت مستويات صعوبتها فى الفترة بين التعديلين نتيجة للتغيرات الثقافية ، وفى أعداد طبعة عام ١٩٦٠ واجه ترمان مشكلة شائعة فى المقياس النفسى وتتلخص (٤٩) فى أنه من المستحب عامة ادخال تعديلات دورية على المقاييس للاستفادة من التقدم العلمى فى ميدان بناء الاختبارات ، ومن الخبرة السابقة باستخدام الاختبار، بالإضافة الى تحديث محتوى الاختبار على نحو يجعله مسايرا لروح العصر . وتزداد الحاجة الى الاعتبار الأخير بالنسبة لأسئلة المعلومات والمواد المصورة التى تتغير تبعا لتغير الظروف الثقافية . فاستخدام أسئلة قديمة يؤثر

فى الصدق الظاهرى للاختبار دون شك ، كما قد يغير فى مستوى صعوبة الأسئلة . ومن ناحية اخرى فان التعديل قد يجعل معظم البيانات المتوافرة من البحوث السابقة على الاختبار لا ترتبط بالصورة الجديدة ، فالاختبارات التى استخدمت استخدامات واسعة لمدة سنوات ، كمقياس ستانفورد - بينيه ، تتوافر عنها ثروة من المعلومات يجب أن توزن فى مقابل الحاجة الى التعديل . ولهذه الاسباب فضل تلاميذ ترمان الجمع بين الصورتين فى صورة واحدة بعد انتقاء افضل ما فيهما . وفقدان صورة مكافئة ليس ثمنا فادحا لتحقيق هذه الغاية ، وخاصة انه أصبحت توجد منذ عام ١٩٦٠ اختبارات فردية اخرى جيدة الصنع لم تكن متوافرة عام ١٩٢٧ .

وفى عام ١٩٧٢ أجرى تقنين جديد للصورة م - ل بإشراف روبرت ثورنديك (٨٠) دون ادخال تعديلات تذكر على محتوى الاختبار ، أما التعديل الجوهرى فقد طرأ على المعايير التى اشتقت من أداء عينة جديدة بلغت ٢١٠٠ مفحوص تم اختبارهم خلال العام الدراسى ١٩٧١ - ١٩٧٢ . وحتى تكون لهذه العينة خصائص التمثيل رغم الصعوبة العملية فى تطبيق الاختبارات الفردية لجأ ناشرو الاختبار الى عينة أكبر بلغت مائتى ألف تلميذ وطالب من الصفوف من الثالث حتى الثانى عشر (فى النظام التعليمى الأمريكى) استخدمت فى تقنين أحد الاختبارات الجماعية المشهورة (اختبار القدرات المعرفية الذى سنشير اليه فى الفصل الثامن) ومثلت خصائص المجتمع الأمريكى من حيث الحجم والمناطق الجغرافية والمستويات الاقتصادية الاجتماعية والأصول العنصرية .

وفى اختيار عينة تقنين مقياس ستانفورد - بينيه من بين هذه العينة الأكبر اعتمد على درجات المفحوصين فى البطارية اللفظية فى اختبار القدرات المعرفية المشار اليه بحيث يصبح توزيع درجات العينة الجديدة مطابقا للتوزيع العام على مستوى العينة الكلية . ولم يستبعد الا اولئك الذين لا يتحدثون اللغة الانجليزية . وحتى تشمل العينة الأعمار من ٢ - ٨ سنوات اختار الباحثون اخوة المفحوصين الذين تضمنتهم العينة وبالمثل اختبرت بعض الحالات الاضافية فى الأعمار العليا . وهكذا أصبحت عينة تقنين المقياس مؤلفة من مائة مفحوص تقريبا لكل فئة عمرية مداها نصف سنة للأعمار من

٢ - ١٣ سنة ، ومائة مفحوص تقريبا أيضا لكل فئة عمرية مداها سنة للاعمار من ٦ - ١٨ سنة .

وفى عام ١٩٨٥ صدرت الطبعة الرابعة من المقياس من أعداد روبرت ثورنديك واليزابيث هاجن وجيروم ساتلر (٨١) . وقد حافظت هذه الطبعة الجديدة على المعالم التاريخية للمقياس التى استمرت فى الطبقات السابقة ، فأبقيت الأسئلة الصالحة من طبعتي ١٩٢٧ ، ١٩٦٠ ، وأضيفت أسئلة أخرى جعلته مقياسا حديثا من حيث المحتوى والخصائص الاحصائية واجراءات التطبيق . واستخدمت فى المقياس الجديد أحدث النماذج السيكمترية وخاصة نظرية السمات الكامنة (أو نظرية الاستجابة على المفردة item response theory ، وحل محل تحيز الأسئلة سواء من حيث الصياغة أو من حيث الخصائص الاحصائية .

ولعل أهم التعديلات التى أدخلت على المقياس فى هذه الطبعة استخدام مفاهيم المقياس التكيفى للقدرة adaptive ability وذلك بتطبيق نظام الدرجات المتعددة وليس الدرجة الواحدة كما كان الحال فى هذا المقياس طوال تاريخه وأمكن بهذا الطريقة اعطاء المفحوص أربعة أنواع من الدرجات خلال زمن الاختبار هى : الاستدلال اللفظى ، والاستدلال الكمي ، والاستدلال المجرد فى مقابل الاستدلال البصرى ، والذاكرة قصيرة المدى ، وبالإضافة الى ذلك يمكن أن يعطى المفحوص درجة كلية مركبة فى المقياس .

وعلى الرغم من أن كل درجة من هذه الدرجات تعكس جزئيا أحد عوامل القدرة العقلية العامة المشتركة فى جميع القدرات ، وعلى الرغم أيضا من الدرجة الكلية المركبة من هذه الدرجات الأربع تعطى تقديرا ثابتا للعامل العام ، فإن كل درجة منها لها معناها السيكلوجى الخاص والذى له أهميته فى ذاته . وقد حسبت المعايير لكل درجة من الدرجات الأربع وللدرجة الكلية . وقد نظم المقياس تبعا لمستويات تبدأ من الطفل الأقل من المتوسط من سن سنتين وحتى مستوى الراشد المتفوق .

تحليل أسئلة المقياس :

لقد استغرق تعديل عام ١٩٢٧ لمقياس ستانفورد - بينيه أكثر من عشر

سنوات من البحث العلمى المنظم حول تجميع الأسئلة الجيدة وتجريبها على جماعات صغيرة العدد من المفحوصين الذين طبق عليهم مقياس عام ١٩١٦ . وكانت الخطوة الأساسية فى ذلك هى تطبيق الصورتين المبدئيتين فى ذلك الوقت على عينة التقنين التى أشرنا إليها (٢١٨٤ مفحوصا) ومنها تم الانتقاء النهائى للأسئلة وتوزيعها على الأعمار وقد أعطيت لكل مفحوص صورتنا الاختبار بفواصل زمنى يتراوح بين يوم واحد وأسبوع واحد . وفى هذه الخطوة كان محك تحليل الأسئلة لتحديد صدقها فى العصر الزمنى والدرجة الكلية المركبة من الصورتين المبدئيتين ، ويرتبط المحك الثانى بما أسميناه الاتساق الداخلى حتى يتحقق فى المقياس درجة من التجانس . واستخدمت ثلاثة مقاييس فى تحليل كل سؤال فى المقياس هى :

١ - منحنيات النسب المثوية للمفحوصين الذين يجتازون السؤال بنجاح فى الأعمار الزمنية المتتابعة .

٢ - منحنيات النسب المثوية للمفحوصين الذين يجتازون السؤال بنجاح فى الفئات المتتابعة للدرجة الكلية فى المقياس (وبالطبع لكل صورة من صورتى المقياس منحنياتها الخاصة) .

٣ - معامل الارتباط الثنائى بين كل سؤال والدرجة الكلية فى المقياس . وتتفق الطريقة الأولى مع أسلوب بينيه الذى كان يستخدمه فى تحليل المفردات وانتقاء الأسئلة التى تصلح للمستويات العمرية المختلفة . وفى هذا يجب أن نشير الى أن ترمان استخدم للحكم على صلاحية الأسئلة نسباً مثوية فى الأعمار الصغيرة أعلى منها فى الأعمار الكبيرة . ومن ذلك مثلاً انه كان يعتبر السؤال مناسباً لسن ٣ سنوات اذا أجاب عليه ٧٢٪ من أطفال هذا السن اجابة صحيحة ، أما فى حالة الأطفال من سن ١٠ سنوات فكانت هذه النسبة ٥٩٪ . ويشير ماكنمر (٦٩) الى أن هذا الاختلاف فى النسب المثوية هو من المطالب الضرورية فى مقاييس العمر . فمن الملاحظ أن الانحراف المعيارى للعمر العقلى يجب أن يتزايد مع العمر اذا كان لا بد لنسبة الذكاء أن تظل ثابتة . وعلى ذلك فاذا كان التباين فى العمر العقلى فى الأعمار العليا أكبر منه فى الأعمار الدنيا فان ذلك يعنى ضرورة وجود انتشار أكبر فى أداء المفحوصين الكبار فى مستويات الأعمار المتجاورة . وعلى ذلك

فمع الزيادة في العمر يقل عدد المفحوصين الذين يجتازون بتجاح السؤال المطابق لمستواهم العمري .

ويمكن أن نوضح هذه المسألة بمثال بسيط لنفرض أنه بين الأطفال من سن ٥ سنوات يصل ٤٠٪ منهم إلى العمر العقلي ٥ ، بينما يصل ٣٠٪ إلى أعمار أعلى ، و ٣٠٪ آخرون إلى أعمار أقل من هذا المستوى . وحيث أن أي مفحوص عمره العقلي ٥ سنوات أو أكبر يجب أن يجتاز من الوجهة النظرية السؤال الخاص بسن ٥ سنوات فإن ٧٠٪ (٤٠٪ + ٣٠٪) من الأطفال من فئة سن ٥ سنوات تتوقع نجاحهم في أسئلة سن ٥ سنوات . ولنفرض من ناحية أخرى أنه من بين الأطفال من سن ١٠ سنوات لا يصل إلى العمر العقلي ١٠ سنوات إلا ٢٠٪ ، بينما يصل إلى أعمار أعلى من ذلك ٤٠٪ منهم ، وإلى أعمار أقل من ذلك ٤٠٪ آخرون . وبالمطبع لا يمكن الحصول على مثل هذا التوزيع إلا إذا كان التباين في الأعمار العقلية لأطفال سن ١٠ سنوات أكبر منه في سن ٥ سنوات . ويعنى ذلك أنه من بين الأطفال من سن ١٠ سنوات تصبح النسبة المتوقعة لاجتياز الأسئلة المخصصة لهذا السن هي (٢٠٪ + ٤٠٪ = ٦٠٪) وفي هذا المثال تصبح النسبة المثوية التي يمكن استخدامها محكا لانتقاء الأسئلة في سن ٥ سنوات هي ٧٠٪ ، وفي سن ١٠ سنوات هي ٦٠٪ .

وفي الممارسة الواقعية لانتقاء أسئلة مقياس ستانفورد - بينيه نجد أن المحك يتناقص من ٧٧٪ عند مستوى سن عامين إلى ما هو أقل قليلا من ٥٠٪ في مستوى الرشد العادي . وتقل النسب عن هذا في انتقاء أسئلة مستويات الرشد المتفوق وذلك للوصول إلى سقف مناسب للاختبار . وعند الانتقاء النهائي للمفردات لا يقتصر الاهتمام على تمايز العمر والاتساق الداخلي فحسب ، وإنما امتد ليشمل أيضا مدى تناقص الفروق بين الجنسين وتوزانها في النسب المثوية للناجحين في مختلف المستويات العمرية . واستبعدت بالفعل المفردات التي اختلف فيها أداء أحد الجنسين عن الآخر بنسب لها دلالة احصائية : على أساس أن مثل هذه المفردات قد تدل على تباين اصطناعي أو غير مرتبط بما يقيسه الاختبار نتيجة للخبرات السابقة للجنسين .

وفى اعداد طبعة عام ١٩٦٠ من المقياس اختبرت الأسئلة من الصورتين ل . م على اساس أداء ٤٤٩٨ مفحوصا يمتد أعمارهم بين ٢ ١/٢ سنة ، ١٨ سنة اجابوا على احدى صورتى المقياس او كليهما فى الفترة من ١٩٥٠ وحتى ١٩٥٤ . كما تضمنت العينة مجموعتين طبقيتين اشتملتا على ١٠٠ طفل من سن ٦ سنوات صُنفت حسب مهنة الوالد ، ١٠٠ مراهق من سن ١٥ سنة صُنفت حسب مهنة الوالد وتوزيع المستوى التعليمى .

ويعنى هذا أن طبعة ١٩٦٠ لم تتضمن اعادة لتقنين المقياس او اعادة لبناء معاييرهِ . واستخدمت العينات الجديدة لمراجعة التغيرات التى طرأت على صعوبة الأسئلة فقط بعد ما تحددت مستويات الصعوبة بإيجاد النسب المثوية للأطفال الذين يجتازون السؤال فى أعمار عقلية متتابعة فى صورتى عام ١٩٢٧ وهكذا فانه بالنسبة لتحليل المفردات تم تجميع المفحوصين تبعاً للعمر العقلى الذى تحدده طبعة عام ١٩٢٧ وليس تبعاً للعمر الزمنى وعلى هذا فان الأعمار العقلية ونسب الذكاء فى طبعة ١٩٦٠ ظلت على النحو الذى كانت عليه فى طبعة ١٩٢٧ وفى ضوء أداء عينات تقنين هذه الطبعة . صحيح أنه روجعت بعض الأسئلة للتحقق من احتمال وجود فروق جغرافية او اقتصادية اجتماعية وذلك من خلال المقارنة بين المجموعات الفرعية التى تكون العينة الكلية . الا ان ما يجب أن ننبه اليه أن المقياس لم يتضمن أى مواد جديدة ، وانما عدلت بعض الرسوم الخاصة بالأشياء الشائعة مسـايرة للعصر . وفى هذه الأحوال اختبرت الأسئلة المعدلة قبلياً على عينات صغيرة قبل تضمينها فى المقياس . وقد ظلت هذه الاستراتيجية فى انتقاء المفردات ثابتة فى طبعتي ١٩٧٢ ، ١٩٨٥ .

وصف المقياس :

يتألف مقياس ستانفورد - بينيه من صندوق به مجموعة من اللعب المقتنة تستخدم مع الأعمار الدنيا ، وكتيبين من البطاقات المطبوعة ، وكراسة لتسجيل الاجابات ، وكراسة التعليمات ، وكراسة معايير التصحيح .

وقد رتبنا أسئلة المقياس وجمعت تبعاً لمستويات الأعمار التى تعتمد من سن عامين حتى سن الرشد المتفوق . وتحدد مستويات المقياس من سن ٢

حتى سن ٥ سنوات فى فئات نصف سنوية على النحو الآتى : ٢ ، ٢ - ٦ ، ٣ ، ٣ - ٦ ، ٤ ، ٤ - ٦ ، ٥ (مع ملاحظة ان الرمز ٢ - ٦ ، ٣ - ٦ ، ٤ - ٦ ، يدل على سن عامين و ٦ شهور ، ٢ سنوات و ٦ شهور ، ٤ سنوات و ٦ شهور على التوالى) وذلك بسبب التقدم العقلى السريع فى الأعمار الدنيا . أما الفترات من سن ٥ سنوات حتى ١٤ سنة فتتحدد مستويات الأعمار فيها فى فئات سنوية ، وتدل المستويات الأخرى فى المقياس على الراشد المتوسط ، والراشد المتفوق (١ ، ٢ ، ٣) . وكان عدد أسئلة كل مستوى عمرى ٦ ، باستثناء مستوى الراشد المتوسط الذى يبلغ عدد أسئلته ٨ .

وكانت الأسئلة فى كل مستوى عمرى من مستوى صعوبة موحدة تقريبا ، وترتب بصرف النظر عن أى فروق طفيفة فى الصعوبة . وقد أعد سؤال احتياطى فى كل مستوى عمرى معادل فى صعوبته لأسئلة العمر . يستخدم هذا السؤال الاحتياطى اذا تطلب الأمر استبعاد أحد الأسئلة العادية بسبب بعض الظروف الخاصة تجعله غير ملائم للمفحوص ، أو بسبب عدم الالتزام فى التطبيق بحرفية التعليمات بما يتعارض مع تقنين المقياس وهكذا يبلغ العدد الكلى لاختبارات المقياس ١٤٢ اختبارا (١٢٢ اختبارا أصليا ، ٢٠ اختبارا احتياطيا) .

ومعنى هذا أنه ابتداء من طبعة ١٩٦٠ من المقياس نجد سؤالا احتياطيا لكل مستوى عمرى ، وهذا مالم يتوافر فى طبعة ١٩٣٧ التى نقلت الى اللغة العربية حيث اقتصر وجود الأسئلة الاحتياطية على المستويات العمرية من سن ٢ حتى سن ٥ سنوات فقط .

وقد اختير ٤ أسئلة من كل مستوى عمرى على أساس الصدق والتمثيل ليتألف منها المقياس المختصر الذى يستخدم حين لا يسمح الوقت باعطاء المقياس الكلى ، وقد وضعت أمام هذه الاختبارات علامة (*) . وتؤكد المقارنات بين نسب الذكاء كما تتحدد بالمقياس الكلى والمقياس المختصر تطابقا كبيرا بينهما حيث يصل معامل الارتباط بينهما الى مستوى معامل ثبات المقياس الكلى (٥٦ ، ٥٧) .

(التقويم النفسى)

تطبيق المقياس :

يتطلب هذا المقياس - كغيره من اختبارات الذكاء الفردية - فاحصا على درجة كبيرة من المهارة والتدريب لأن تطبيق هذا المقياس وتصحيحة على درجة كبيرة من التعقد ويستلزم أن تتوافر في الفاحص الفة شديدة به وخبرة طويلة في التطبيق . لأن تردد الفاحص اثناء التطبيق قد يؤدي الى عجز عن تكوين علاقة بينه وبين المفحوص ، بل وفقدان ثقة المفحوص به . بل ان أدنى تغيير في صياغة الأسئلة قد يؤدي الى تغيير في مستوى صعوبتها كما تتعدد في عينة التقنين . بالاضافة الى ما يتطلبه المقياس من تصحيح فوري لاختباراته عقب اعطائها مباشرة ، لأن تتابع اعطاء الاختبارات يعتمد على أداء المفحوص في المستويات السابقة .

ويرى كثير من علماء النفس الاكلينيكين ان مقياس ستانفورد - بينيه ليس مقياسا مقننا للذكاء فحسب ، وانما هو كذلك نوع من المقابلة الاكلينيكية ، لأنه يهيئ فرصة للتفاعل بين الفاحص والمفحوص ، ويهيئ مصادر أخرى للدلالات الاكلينيكية لدى الأخصائي الاكلينيكي المتمرس ، بل ان مقياس ستانفورد - بينيه ، أكثر من غيره من الاختبارات الفردية ، يهيئ الفرصة للفاحص أن يلاحظ الطرق التي يلجأ اليها المفحوص في العمل ، وأساليب حله للمشكلات ، وغير ذلك من الأساليب الكيفية للأداء ، بل ان الفاحص يمكنه ان يحكم من خلال الموقف الاختباري الذي يستخدم فيه هذا المقياس على بعض خصائص الشخصية مثل مستوى النشاط والثقة بالنفس والمثابرة والقدرة على التركيز . وبالطبع فان أي ملاحظات كيفية يمكن تسجيلها اثناء موقف الاختبار يجب تحديدها في ذاتها ، ولا يجب تفسيرها على نفس النحو الذي تفسر به درجات الاختبار الموضوعي . وتعتمد قيمة هذه الملاحظات الكيفية الى حد كبير على مهارة الفاحص وخبرته واعداده السيكولوجي العميق ، بالاضافة الى وعيه الشديد بمزالق وحدود هذا النوع من الملاحظات المباشرة . وقد استطاع مورياتري (٧٠) أن يسجل انماط الملاحظات الاكلينيكية في موقف الاختبار الفردي ، على أساس ان جلسة الاختبار هي لدراسة سلوك الطفل اثناء مواجهة موقف تسود فيه عوامل التحدي والاستثارة والصعوبة والاحباط .

وفى اعطاء مقياس ستانفورد - بينيه لا يحاول المفحوص جميع أسئلة المقياس ، وإنما يختبر المفحوص فقط حول مدى المستويات العمرية الملائمة لمستواه العقلى . ولا يتطلب المقياس أكثر من ٢٠ - ٤٠ دقيقة للأطفال الصغار ، ولا يتطلب أكثر من ساعة ونصف للراشدين . والاجراء المقتن فى تطبيق المقياس هو أن تبدأ الاختبار من مستوى أدنى قليلا من العمر العقلى المتوقع للمفحوص وعلى ذلك فان الاختبارات الأولى يجب أن تكون سهلة بالقدر الذى يستثير فى المفحوص الثقة بالنفس ، بشرط ألا تكون سهلة جدا بحيث تؤدي بالمفحوص الى الاستهانة بالاختبار أو الملل . والاجراء الأكثر موضوعية فى هذا الصدد هو أن تبدأ باعطاء المفحوص أسئلة العمر الأدنى مباشرة من عمره الزمنى ، فإذا فشل فى أى سؤال من أسئلة هذا المستوى يعطى أسئلة المستوى الأدنى منه ، وهكذا يستمر الفاحص حتى يصل الى العمر الذى يستطيع المفحوص أن يجيب على جميع أسئلته ، ويصبح هذا المستوى هو « العمر القاعدى » basal age وبعد ذلك نعطيه أسئلة عمره الزمنى فان أجاب عليها جميعا أو على بعضها تعطيه أسئلة الأعمار الأعلى منه تدريجيا حتى نصل الى العمر الذى يفشل المفحوص فى الاجابة على جميع أسئلته ويسمى هذا المستوى « الحد الأعلى للعمر » أو سقف الاختبار ceiling age . وحين يصل الفاحص الى هذا المستوى يتوقف عن اعطاء المقياس .

تصحيح المقياس :

تصحح جميع أسئلة مقياس ستانفورد - بينيه على اساس النجاح أو الفشل (واحد أو صفر) . وتحدد كراسة التعليمات الأداء المطلوب « للنجاح » فى كل سؤال . وقد تظهر نفس الأسئلة فى مستويات عمرية مختلفة ولكنها تصحح بمستويات مختلفة للنجاح . وفى هذه الحالة تطبق الأسئلة مرة واحدة ثم يحدد أداء المفحوص المستوى العمرى الذى يستحقه . ومن ذلك مثلا أن سؤال المفردات يطبق ابتداء من المستوى العمرى ٦ سنوات حتى المستوى الثالث للراشد المتفوق اعتمادا على عدد الكلمات التى يستطيع المفحوص تحديدها وتعريفها .

والواقع أن الأسئلة التي ينجح فيها المفحوص أو يفشل تظهر مقدار
من التشنت في المستويات العمرية المتتابة . فنحن لا نجد مفحوصا ينجح
على جميع أسئلة عمره العقلي أو أدنى منه ويفشل في جميع الأسئلة الأعلى
من هذا المستوى ، وإنما نجد في العادة أن الأسئلة التي ينجح فيها تنتشر
في مستويات عمرية عديدة يحددها العمر القاعدي من ناحية والحد الأعلى
للعمر (سقف الاختبار) من ناحية أخرى . وبحسب العمر العقلي في هذه
الحالة بالعمر القاعدي مضافا إليه أي شهور عقلية لأي اختبار اجتازه من
مستويات أعلى من عمره القاعدي تبعا للأوزان المبينة في الجدول رقم (٨) .

جدول رقم (٨) يبين أوزان أسئلة مقياس ستانفورد - بينيه في
الأعمار المختلفة

العمر	عدد الأسئلة	وزن السؤال مجموع التقدير الوزني بالشهر لكل مستوى بالشهر
٢	٦	٦
٢ - ٦	٦	٦
٣	٦	٦
٣ - ٦	٦	٦
٤	٦	٦
٤ - ٦	٦	٦
٥	٦	٦
٦	٦	١٢
٧	٦	١٢
٨	٦	١٢
٩	٦	١٢
١٠	٦	١٢
١١	٦	١٢
١٢	٦	١٢
١٣	٦	١٢
١٤	٦	١٢
الراشد المتوسط	٨	١٦
الراشد المتفوق (١)	٦	٢٤
الراشد المتفوق (٢)	٦	٣٠
الراشد المتفوق (٣)	٦	٣٦

المجموع	١٢٢	المجموع	٢٥٦
+	+		
عدد الأسئلة الاحتياطية	٢٠	العمر الأدنى لبدء الاختبار	١٨
المجموع الكلى	١٤٢	شهور	٢٧٤

سنة
٢٢
شهور
١٠
أى

ومن الوجهة النظرية فإن أعلى عمر عقلى يمكن الوصول إليه بمقياس ستانفرد - بينيه - هو ٢٢ سنة و ١٠ شهور . والواقع أن هذه الدرجة لاتدل على عمر عقلى حقيقى ، وإنما هى درجة عددية تدل على مقدار الزيادة عن الأداء المتوسط الذى يصدر عن الراشد العادى . وهى بالتأكيد لا تطابق ما يحصل إليه الشخص المتوسط من عمر زمنى مقداره ٢٢ سنة و ١٠ شهور ، لأن هذا الشخص العادى يحصل على عمر عقلى مقداره ١٦ سنة و ٨ شهور كما تحدده معايير ١٩٧٢ (*) . أما بالنسبة لأى راشد يزيد عمره عن ١٨ سنة فإن أى عمر عقلى مقداره ١٦ سنة ، ٨ شهور يحصل على نسبة الذكاء ١٠٠ .

ويجب أن نؤكد هنا أنه بالنسبة للمفحوصين ابتداء من مرحلة المراهقة (١٢ +) لا يعطى العمر العقلى فى هذا المقياس نفس الدلالة التى يعطيها فى الأعمار الأدنى ، لأنه بعد سن ١٢ يقل متوسط الأعمار العقلية عن متوسط الأعمار الزمنية ، ولذلك فإن مقياس ستانفرد - بينيه لا يلائم الراشدين وبخاصة العاديين والمتفوقين منهم . ورغم وجود ٢ مستويات للراشد المتفوق فيه فإن « سقف » الاختبار لا يلائم معظم الراشدين المتفوقين أو حتى الراشدين العاديين كما تؤكد دراسات كيندى وزملائه (٦٧) فمن المستحيل أن نجد فى المقياس سقفا مناسباً لهؤلاء ، أى مستوى عمرى يعجز المفحوصون فى الإجابة عن جميع أسئلته وبالإضافة الى ذلك فإن مقياس ستانفرد - بينيه أكثر جاذبية للصغار منه للمراهقين والراشدين من حيث محتواه .

ومن أهم الإضافات الى طبعة عام ١٩٦٠ احلال نسبة الذكاء الانحرافية محل نسبة الذكاء العادية التى استخدمت فى المقياس منذ ظهوره عام ١٩١٩ . وهذه النسبة الجديدة عبارة عن درجة معيارية معسدة متوسطة لها ١٠٠ وانحرافها المعيارى ١٦ . والميزة الأساسية فى هذا النوع من المعايير أنها تعطى درجات قابلة للمقارنة فى مختلف الأعمار . وبهذا أمكن التغلب على مشكلة عدم استقرار نسبة الذكاء العادية . فعلى الرغم من الجهود التى بذلت عام ١٩٢٧ للحصول على تشتتات ثابتة لنسبة الذكاء فى مختلف الأعمار ، فإننا نلاحظ أن هذا المقياس تذبذبت فيه الانحرافات المعيارية

(*) كان هذا العمر ١٦ سنة فى طبعة ١٩٢٧ .

لمختلف الأعمار من انحراف معيارى ١٢ (فى سن ٦) الى انحراف معيارى ٢١ (فى سن ٢ - ٦) ومعنى هذا أن نسبة الذكاء ١١٢ فى سن ٦ تقابل نسبة الذكاء ١٢١ فى سن ٢ - ٦ . وقد وضع مكنمار جداول خاصة بتصحيح آثار هذه الاختلافات فى طبعة عام ١٩٢٧ (فى ٧٨) وتتلخص فكرة التصحيح فى أنه إذا كان العمر العقلى للمفحوص بين سن ١٢ ، ١٦ فإنه يجب تعديل العمر الزمنى الذى يؤخذ فى الاعتبار عند حساب نسبة الذكاء ، وذلك بأن نضيف ثلثى الشهور الإضافية التى عاشها المفحوص بعد سن ١٢ . فمثلا إذا كان العمر الزمنى الحقيقى ١٤ سنة فإنه يحسب ١٢ - ٨ ، وإذا كان ١٥ سنة يحسب ١٤ - ٤ .

وقد أمكن التغلب على كل هذه الصعوبات فى طبعة ١٩٦٠ باستخدام نسبة الذكاء الانحرافية التى تعنى استخدام انحراف معيارى موحد لجميع الأعمار . وقد أعد بينو (٧٢) جداول تستخرج منها مباشرة نسبة الذكاء الانحرافية من كل من العمر العقلى والعمر الزمنى بالسنوات والشهور . وقد نشرت هذه الجداول فى كراسة تعليمات المقياس الصادرة عام ١٩٧٢ (٨٠) .

وتعتمد فكرة التحويل فى جداول بينو على اعتبار نسبة الذكاء التقليدية هى الدرجة الخام للمفحوص تحول الى نسبة ذكاء انحرافية باستخدام متوسطات وانحرافات معيارية تجريبية لنسب الذكاء التقليدية لكل فئة من الفئات العمرية ، ثم يستخدم الانحراف المعيارى الجديد ١٦ والمتوسط الجديد ١٠٠ ، أى أن معادلة نسبة الذكاء الانحرافية فى مقياس ستانفورد - بينيه هى:

نسبة الذكاء الانحرافية =

$$100 + \left[16 \times \left(\frac{\text{نسبة الذكاء التقليدية} - \text{المتوسط التجريبي}}{\text{الانحراف المعياري التجريبي}} \right) \right]$$

$$\text{أو} \left[\frac{16}{\text{الانحراف المعياري}} \times (\text{نسبة الذكاء التقليدية} - \text{المتوسط التجريبي}) \right] + 100$$

وقام بينو بقسمة ١٦ على الانحراف المعياري التجريبي لكل فئة عمرية
فحصل على ما اسماه بالمقدار الثابت K وأعد جدولاً بهذه الثوابت
لفئات عمرية مداها شهر . والجدول رقم (٩) يوضح هذه الثوابت لفئات
العمر المختلفة إلا أن ما يجب أن نشير إليه هو أن حساب العمر العقلي
لا يزال يتم بالطريقة المستخدمة منذ عام ١٩٣٧ . أي أن التعديلات أدخلت
على مفهوم نسبة الذكاء التقليدية دون أن تمس مفهوم العمر العقلي ذاته .

جدول (٩) يبين الثوابت اللازمة لتحويل نسبة الذكاء التقليدية الى
نسبة ذكاء انحرافية .

العمر الزمني	المتوسط التجريبي	ك	العمر الزمني	المتوسط التجريبي	ك
٢	١٠٢	٩٩ر	٨٠	١٠٢	١٠٢ر
٢-٦	١٠٣	٧ر	٩	١٠٣	٩٨ر
٣	١٠٤	٨٨ر	١٠	١٠٣	٩٤ر
٣-٦	١٠٢	٩٤ر	١١	١٠٢	٩١ر
٤	١٠٢	٩٨ر	١٢	١٠٢	٨٨ر
٤-٦	١٠١	١٠٤ر	١٣	١٠١	٨٩ر
٥	١٠٠	١٠٩ر	١٤	١٠١	٩٠ر
٦	١٠٠	١١١ر	الراشد	١٠٥	٩٠ر
٧	١٠١	١٠٧ر			

ويوجد تعديل آخر اضافته طبعة ١٩٦٠ واكدته طبعة ١٩٧٢ . واعتمد
على النتائج التي تؤكد أن التحسن في الأداء في المقياس يستمر الى سن
١٨ سنة بدلاً من التوقف عند سن ١٦ كما كان الحال في طبعة
١٩٣٧ . فقد بنيت البحوث الطولية في النمو العقلي التي أجريت خلال الربع
قرن بين طبعتي المقياس أن القدرات العقلية كما تقاس باختبارات الذكاء

تستمر في التحسن والارتقاء أبعد مما كان مفترضا (٢٤) . وأكثر الأدلة مباشرة على زيادة التحسن في مقياس ستانفرد - بينيه بعد سن ١٦ قدمته بحوث برادوى وطومسون وكرافتر (٨٢) ، بعد انقضاء ١٠ سنوات ثم بعد انقضاء ٢٥ سنة ، ف لوحظ أن هؤلاء المفحوصين حين اختبروا أول الأمر وكانت أعمارهم الزمنية ٢ - ٥ ½ سنة ثم اختبروا بعد ١٠ سنوات ، أن متوسطهم في نسبة الذكاء لم يختلف في الحالتين عما حددته معايير المقياس ولكن هذا المتوسط أظهر ارتفاعا له دلالة احصائية بمقدار ١١ر٣ نقطة بين الاختبار الثانى والثالث (أى بعد انقضاء ١٥ سنة من الاختبار الثانى و ٢٥ سنة من الاختبار الأول) . وهذه الزيادة نتجت عن أن المفحوصين استمروا في التحسن بعد سن ١٦ سنة ، بينما نجد أن حساب نسب الذكاء الذى اعتمدت عليه طبعة ١٩٢٧ افترض توقف النمر العقلى عند هذا السن .

وعند المقارنة بين طبعات ١٩٢٧ ، ١٩٧٢ ، ١٩٨٥ نجد أن المعايير في الطبعات الأحدث تعتمد على عينات أكثر تمثيلا ، كما أنها أكثر حداثة وتعكس أى تغيرات في الظروف الحضارية والثقافية يمكن أن تؤثر في الأداء الاختباري . ومن الطريف أن نذكر أن المعايير الأحدث تظهر بعض التحسن في أداء الاختبار في جميع الأعمار ، ويكون هذا التحسن أكثر جوهرياً في أعمار ما قبل المدرسة يصل في متوسطه الى حوالى عشر نقاط . ويرجع مؤلفو الاختبار هذا التحسن الى أثر وسائل الاعلام من ناحية وإلى ارتفاع المستوى الثقافى والتعليمى للوالدين .

طريقة حساب العمر العقلى ونسبة الذكاء التقليدية :

يمكن توضيح الأسلوب الذى يستخدم في الحصول على العمر العقلى ونسبة الذكاء التقليدية من مقياس ستانفرد - بينيه بالمثال الموضح في الجدول رقم (١٠) .

جدول رقم (١٠) يوضح طريقة حساب العمر العقلي ونسبة
الذكاء التقليدية

العمر	عدد الأسئلة التي أجاب عليها المفحوص	وزن السؤال بالشهر	المجموع الكلي للأوزان
أقل عمره الزمني ٦ سنوات ، ٤ شهور شهر سنة			
٤	٦	العمر القاعدي	٢
٤ - ٦	٥	١	٥
٥	٣	١	٢
٦	٣	٢	٦
٧	٢	٢	٤
٨	١	٢	٢
٩	صفر	سقف الاختبار	-
			٤
العمر العقلي = ٨	سنة	نسبة الذكاء	سنة
العمر الزمني = ٤	٦	التقليدية = ٨٩	

وبالطبع تحول نسبة الذكاء التقليدية ٨٩ الى نسبة ذكاء انحرافية
باستخدام جداول بينو المشار اليها .
تفسير معايير المقياس :

من أهم مزايا مقياس ستانفورد - بينيه ما يترافق حصوله من بيانات
امبريقية وخبرة كلينيكية تراكمت طوال ثمانين عاما منذ مقياس بينيه المبكر .
وقد أصبح هذا المقياس عند كثير من علماء النفس الكلينيين والتربويين
وغيرهم من المهمتين بالتقويم النفسى مرادفا للذكاء . وأصبح تحديد مستويات
الذكاء يعتمد فى جوهره على نسب الذكاء كما يحددها المقياس ، رغم ما

فى هذا الرأى من غلو لا يتسع المقام لمناقشته . والجدول رقم (١١) يوضح التوزيع والتصنيف النظريين لنسب الذكاء الانحرافية لطبعتى ١٩٦٠ ، ١٩٧٢ .

جدول رقم (١١) يبين التوزيع والتصنيف النظرى لنسب الذكاء الانحرافية لمقياس ستانفورد - بينيه

الدرجة الميارية	نسب الذكاء الانحرافية	التصنيف	النسب المئوية للحالات
٣ -	٥٢	التخلف العتى (الضعف العقلى)	١.٢٪
٢ -	٦٤	الحالات الهامشية (قد تصنف أغبياء أو ضماف عقول	٥.٥٪
١ -	٦٨	أو ضماف عقول	
	٧٦		
	٨٤	الغباء (بطء التعلم)	١.٦٪
	٨٨		
صفر	١٠٠	المتوسط (أو العادى)	٥٤.٧٪
١ +	١١٢		
	١١٦	المتفوقون	١.٦٪
	١٢٤		
٢ +	١٣٢	المتفوقون جداً	٠.٦٪
٣ +	١٤٨		
		العبقرية أو شبه العبقرية	٠.١٪
		أعلى من ذلك	

وقد كان من المعتاد لعدة سنوات استخدام نسبة الذكاء التقليدية ٧٠ باعتبارها نقطة قطع تقريبية للتخلف العقلى أو الضعف العقلى المحتمل ، وتتطابق هذه الفئة مع النسبة ١٪ أو ٢٪ من الحد الأدنى للتوزيع . أما الجدول فيوضح لنا أن نقطة القطع هذه أصبحت عند الدرجة المعيارية (- ٢٢٨) والتي تتطابق مع نسبة الذكاء الانحرافية ٦٤ والتي يقع أدنى منها ١.٢٪ من الحالات .

وعموما فان الفضل يعود الى هذا المقياس فى استخدام التقسيم الثلاثى الشهير للضعف العقلى : المأفونون والبلهاء والمعتوهون ، وهى الفاظ لم تستخدم فى الوقت الحاضر وحل محلها ما يلى :

١ - الحالات القابلة للتعليم educable وتمتد نسب ذكائها من ٥٠ ، ٧٥ ويمكنها ان تتعلم حتى الصف الثالث فى المرحلة الابتدائية ولا يتجاوز حدما الاقصى فى التعلم نهاية المرحلة الابتدائية اذا تعلمت مرقف تعليمى خاص .

٢ - الحالات القابلة للتدريب trainable وتمتد نسب ذكائها بين ٢٥ ، ٥٠ ويمكنها ان تتعلم رعاية الذات والتوافق الاجتماعى فى بيئة حماية خاصة .

٣ - الحالات التى تحتاج الى الايداع فى المؤسسات custodial وتقل نسب ذكائها عن ٢٥ وتتطلب رعاية وحماية خاصة .

والتصنيف الراهن للضعف العقلى يقسم هذه الحالات الى اربع فئات يوضحها الجدول رقم (١٢) .

جدول رقم (١٢) يبين مستويات الضعف العقلى فى صورة نسب ذكاء انحرافية

المستوى	(وحدات ع)	الذكاء والانحراف النسب المئوية مدى نسب (ع = ١٦) للحالات	نقاط القطع
الضعف العقلى الخفيف	٢ -	٥٢ - ٦٨	٢١٤
الضعف العقلى المتوسط	٣ -	٢٦ - ٥١	١٣
الضعف العقلى الشديد	٤ -	٢٠ - ٣٥	٠٠٣
الضعف العقلى الخطير	٥ -	١٩ فأدنى	٠٠٠٠٣

والميزة الجوهرية فى هذا التصنيف انه يصلح للتحويل الى درجات معيارية أو نسب ذكاء انحرافية فى اى مقياس للذكاء .

والواقع أن شيوع تصنيف مستويات نسب الذكاء - رغم فائدته فى تقنين تفسير الأداء الاختبارى - يتضمن بعض المخاطر . فهو كغيره من صور التصنيف السلوكى يجب أن يستخدم بعرونة كافية ، كما يجب ألا يستبعد البيانات الأخرى التى تتوافر عن المفحوص . وبالطبع لا يوجد حد فاصل تماما بين فئات التصنيف . ومن ذلك مثلا لا يوجد خط يميز تماما بين التخلف العقلى والحالات الهامشية ، أو بين التفوق والتفوق الشديد . فقد لوحظ مثلا أن بعض المفحوصين من ذوى نسب الذكاء ٦٠ يتكيفون تكيفا عاديا لمطالب الحياة اليومية ، بينما نجد بعض الأشخاص من ذوى نسب الذكاء ١٠٠ يتطلبون الإبداع فى المؤسسات ، كما أن بعض الأشخاص من ذوى نسب الذكاء ١٦٠ قد لا يعيشون حياة متميزة ، بينما نجد البعض الآخر من ذوى نسب الذكاء الأقرب إلى ١٠٠ يسهمون بعض الاسهامات الممتازة . ومعنى ذلك أن القرارات التى تتحدد حول تشخيص الضعف العقلى أو التفوق العقلى يجب ألا تعتمد على نسب الذكاء وحدها ، وإنما يجب فى حالة الضعف العقلى والقرارات المتصلة به كالإبداع فى المؤسسات أو الالتحاق بمدارس التربية الفكرية الاهتمام بمتغيرات أخرى مثل النضج الاجتماعى والتوافق الانفعالى والظروف الصحية والجسمية والأحوال الاسرية والظروف المدرسية وغير ذلك من العوامل التى تؤثر فى الأداء العقلى وفى جميع الحالات يجب الاهتمام بكل من النشاط العقلى للفرد ونضجه الاجتماعى وسلوكه التوافقى جميعا وفى وقت واحد . وبالمثل فإن نسب الذكاء العالية لا تترادف العبقرية ، لأن الانجاز الرفيع قد يتطلب بالإضافة الى ذلك الابتكار والمواهب الخاصة والمثابرة والاهتمام بالهدف وغيرها من العوامل العقلية والانفعالية والدافعية (٢٢) .

ثبات المقياس :

اجريت دراسات الثبات على طبعة ١٩٢٧ من المقياس باستخدام طريقة النصور المتكافئة على فترات زمنية مقدارها اسبوع أو اقل (للتحقق من مصدرى تباين الخطأ : الاستقرار الرقضى وتكاثر المعترى) . وقد لوحظ

أن المقياس يميل الى أن يكون أكثر ثباتا فى الأعمار الكبيرة منه فى الأعمار الصغيرة ، وفى مستويات نسب الذكاء الدنيا عنه فى نسب الذكاء العليا .

والواقع أن زيادة الثبات مع الزيادة فى العمر هى من خصائص الاختبارات النفسية بوجه عام ، وتنتج جزئيا عن التحكم الأفضل فى ظروف اعطاء الاختبار فى حالة المفحوصين الكبار السن (وبخاصة حين نقارنهم بأطفال ما قبل المدرسة) . بالإضافة الى العوامل الأخرى المسهمة فى زيادة معاملات الثبات ومنها بطء معدل النمو مع التقدم فى العمر .

أما عن ارتفاع معاملات الثبات فى حالة نسب الذكاء الدنيا فى أى عمر زمنى فترتبط فيما يبدو بخصائص مقياس ستانفورد - بينيه على وجه الخصوص . فقد أشرنا الى أن الأوزان النسبية لأسئلة الأعمار المختلفة تختلف ، فهى فى الأعمار الدنيا مقدارها شهر ، وفى الأعمار المتوسطة شهران ، وفى المستويات العليا قد تكون ٤ شهور أو ٥ شهور أو ٦ شهور . ومعنى ذلك أن طريقة وزن الدرجات تميل الى « تضخيم » أخطاء القياس فى المستويات العليا ، لأن النجاح أو الفشل الناجم عن الصدفة فى سؤال واحد يؤدى الى فرق أكبر فى الدرجة الكلية فى هذه المستويات عنه فى المستويات الدنيا . وحيث أن المفحوصين من ذوى نسب الذكاء العالية (من أى عمر زمنى) عادة ما يجيبون على الاسئلة التى تقع فى المستويات العليا من المقياس فان نسب ذكائهم نتيجة لذلك تتضمن مقدارا أكبر من أخطاء القياس وبالتالي تنخفض معاملات الثبات .

وتؤكد هذه الحقيقة من ملاحظة جدول انتشار درجات الصورتين م ، ل من طبعة ١٩٢٧ الذى تؤكد أنه شكل المروحة فى أى عمر وهذا يعنى أن الصورتين ترتبطان عاليا فى المستويات الدنيا من نسب الذكاء ، أما فى المستويات العليا فان العلاقة بين الصورتين تتناقص بوضوح .

وفى مثل هذه الاحوال يكون من المفضل حساب معامل ارتباط كلئى (كمؤشر للثبات) ، ولهذا حسبت معاملات ثبات مختلفة للفتات المختلفة من نسب الذكاء فى كل مرحلة عمرية .

وبوجه عام يمكن القول أن مقياس ستانفورد - بينيه على درجة عالية

من الثبات فى مختلف المستويات العمرية وفى مختلف نسب الذكاء ، بالرغم من تجانس مجموعات عينة التقنيين من حيث العمر الزمنى ، مما يؤثر بالطبع فى معاملات الارتباط عموما ، ويؤدى الى نقصانها ٠٠ وقد تراوحت هذه المعاملات بالنسبة للاعمار $2\frac{1}{4}$ - $5\frac{1}{4}$ بين ٨٢ لذوى نسب الذكاء العالية ، ٩١ ر٠ لذوى نسب الذكاء المنخفضة ٠ وبالنسبة للاعمار من ٦ - ١٣ تتراوح هذه المعاملات بين ٩١ ر٠، ٩٧ ر٠ ولنفس المستويات من نسب الذكاء ، وبالنسبة للاعمار من ١٤ - ١٨ كانت معاملات الثبات المطابقة للمستويين هى ٩٥ ر٠ ، ٩٨ ر٠ .

صدق المقياس :

(١) الصدق المرتبط بالمحككات : تتوافر بيانات كافية عن الصدق المرتبط بالمحككات سواء كان صدقا تلازميا أو صدقا تنبؤيا فى ضوء محك التحصيل المدرسى ٠ فمقد ظهر مقياس مستانفرد - بينيه عام ١٩١٦ لم يتوقف الباحثون عن حساب معاملات الارتباطات بينه وبين الدرجات المدرسية وتقديرات المعلمين ودرجات الاختبارات التحصيلية المقتنة ٠ ويتراوح معظم هذه المعاملات بين ٤٠ ر٠ ، ٧٥ ر٠ كما ارتبط بالتفروق والتخلف الدارسين بوجه عام ٠

وقد وجد أن هذا المقياس - كغيره من مقاييس الذكاء - يرتبط ارتباطا عاليا بالأداء فى جميع المواد الدراسية الأكاديمية تقريبا ، إلا أن ارتباطه الأعلى ما يكون بالمقررات اللغوية (كمواد اللغات أو المواد الاجتماعية) ٠ وقد تأكدت هذه النتائج عند اختيار التلاميذ والطلاب فى مختلف المراحل التعليمية ، وبالطبع فإن معاملات الارتباط تميل الى التناقص مع ارتفاع السلم التعليمى بسبب عوامل الانتقاء وعدم ملائمة سقف الاختبار مما يؤدى الى نقصان التباين بين المفحوصين ٠

وفى تفسير نسبة الذكاء يجب أن نؤكد أن مقياس مستانفرد - بينيه (كغيره من اختبارات الذكاء العام) هو فى الأغلب مقياس للاستعداد المدرسى أو الأكاديمى العام ٠ ومشبع الى حد كبير بالمحتوى اللفظى ، وبخاصة فى اختبارات المستويات العمرية العليا ٠ ولذلك فإن الأفراد من ذوى النقائص

اللغوية ، او أولئك الذين يتفوقون فى بعض القدرات غير اللغوية (كالقدرات الموسيقية) يحصلون على درجات منخفضة نسبيا فى المقياس . وكذلك فانه توجد بعض المجالات لا يلعب فيها الاستعداد المدرسى او الفهم اللغوى دورا هاما ، وفيها لا تصلح اختبارات الذكاء من نعط ستانفورد - بينيه . ومعنى ذلك ان استخدام الاختبارات بصفة عامة يجب ان يكون فى المواقف التى تلائمها ، والا فان الباحث يتوقع من المقياس بما لا يعطيه .

(ب) صدق التكوين الفرضي : من أهم المحكات التى استخدمت فى انتقاء اختبارات مقياس ستانفورد - بينيه محك تمايز العمر . ومعنى ذلك ان المقياس يقيس قدرات متزايدة مع العمر منذ الطفولة حتى الرشد (فى ظل اطار ثقافى وحضارى معين بالطبع) .

وقد استخدم أيضا محك الاتساق الداخلى فى انتقاء الأسئلة . وقد أكدت النتائج وجود نوع من التجانس الوظيفى فى مقياس ستانفورد - بينيه ، بالرغم من التنوع الظاهرى فى المحتوى ، والدليل على هذا ان متوسط معاملات الارتباط بين السؤال والمقياس ككل هو ٠.٦٦ فى طبعة عام ١٩٦٠ . الا أن من الواضح غلبة الطابع اللفظى على المقياس ، ويتضح هذا من ارتفاع معامل الارتباط بين الأسئلة اللفظية والاداء الكلى فى المقياس عنه فى حالة الأسئلة غير اللفظية .

ومن البراهين الاضافية على صدق التكوين الفرضى نتائج الدراسات العاملة العديدة التى أجريت على اختبارات مقياس ستانفورد - بينيه . ويعتمد هذا البرهان على انه لو كانت نسب الذكاء قابلة للمقارنة فى مختلف الأعمار ، فان المقياس يجب ان يكون له نفس التكوين العاملى فى جميع مستويات الأعمار ، بل يجب ان يكون المقياس مشبعا بعامل عام واحد حتى يمكن تفسير نسبة الذكاء تفسيراً واضحاً . أما اذا كان المقياس مشبعا بعاملين طائفيين او اكثر فان نسبة الذكاء الواحدة التى يحصل عليها افراد متعددون (نسبة ذكاء ١٢٠ مثلا) قد لا تدل على معنى سيكولوجى واحد .

وقد تحقق هذا الغرض من نتائج التحليلات العاملية التي أجراها ماكنمر (٦٩) لاختبارات مقياس ستانفورد - بينيه في ١٤ مستوى عمريا ابتداء من سن سنتين وحتى سن ١٨ سنة . وكان عدد المفحوصين في كل تحليل عاملي يتراوح بين ٩٩ ، ٢٠٠ ، كما تراوح عدد الأسئلة التي خضعت للتحليل بين ١٩ ، ٢٥ . وفي كل الأحوال حسبت معاملات الارتباط الرباعية بين المفردات ، وخضعت مصفوفات الارتباط للتحليل العاملي . وتؤكد نتائج ماكنمر أن الأداء في مقياس ستانفورد - بينيه يمكن تفسيره في ضوء عامل عام واحد ، كما وجدت عوامل طائفية إضافية في بعض المستويات العمرية ، إلا أن أسهام هذه العوامل في التباين كان قليلا . كما ثبت أن العامل العام المشترك الذي وجد في المستويات العمرية المتقاربة كان متشابهة إلى حد كبير . ومن أهم النتائج التي توصل إليها ماكنمر أن العامل العام يصبح أكثر لفظية كلما تقدمت الأعمار . ومن ذلك مثلا أن تشبع العامل العام باختبار المفردات ارتفع من ٥٩٪ في سن ٦ سنوات إلى ٩١٪ في سن ١٨ سنة .

وأجرى جونز (٦٥ ، ٦٦) دراستين عامليتين لأسئلة مقياس ستانفورد - بينيه على ٤ مجموعات من الأطفال أعمارهم ٧ ، ٩ ، ١١ ، ١٢ سنة على التوالي ، وكانت كل مجموعة تتكون من ١٠٠ ذكر ، ١٠٠ أنثى ، ووجد أنه في كل عمر يوجد عدد من القدرات ، التمايزة ولكنها مرتبطة ، ومنها العوامل اللفظية والتذكرية الاستدلالية والمكانية والادراكية . وقد اختلفت هذه العوامل وأوزانها النسبية من مستوى عمري لآخر . وتتفق هذه النتيجة مع نتائج دراسات عاملية أخرى على هذا المقياس يمكن الرجوع إليها ملخصة في (١٨) .

والتوقع أن الأساليب الإحصائية التي استخدمها ماكنمر وجونز من النوع الذي يؤكد دور العامل العام في الحالة الأولى (دراسات ماكنمر) ، ودور العوامل الطائفية المختلفة في الحالة الثانية (دراسات جونز) . وحين نتأمل النتائج التي توصل إليها الباحثان نجد أنها تؤكد وجود قدر مشترك في المقياس ككل ، وهي خاصية من خصائص بناء المقياس حسب طريقة انتقاء أسئلته التي كانت تؤكد ارتفاع معاملات الارتباط بين السؤال والدرجة الكلية في المقياس وبالإضافة إلى ذلك فإن الأداء يتأثر كذلك بعدد من القدرات المتخصصة التي يختلف تكوينها تبعاً لمستوى العمر الذي نختبره .

(التقويم النفسي)

أما عن الدراسات التي أجريت في مصر لتحديد المعالم السيكومترية لقياس ستانفورد - بينيه فإنها قليلة . وتتلخص في دراسة قام بها لويس كامل مليكه عام ١٩٦٠ (٣٤) للتحقيق من صدق الترتيب المبدئي لاختبارات المفردات في المقياس ، وذلك بتطبيقها على مجموعات كبيرة من الأفراد ومن مختلف فئات السن (بين ٣ سنوات ، ٢٠ سنة) بلغ عددها ١٦٤٢ مفحوصا ، وفي عام ١٩٦٢ قام محمود الهوارى (٢٦) بدراسة عن ثبات المقياس وصدقه ، استخدم في حساب الثبات طريقة إعادة الاختبار على ٩٥ تلميذا بالسنة الثالثة الاعدادية (متوسط أعمارهم ١٤ عاما) ، بفاصل زمني بين شهر وشهرين ، وبلغ معامل الارتباط ٠.٨٢ . أما عينة الصدق فتألفت من ١٦٠ تلميذا بالسنة الثالثة الاعدادية أيضا ، وقد حسب ليم معاملات الارتباط بين درجات المقياس (نسبة الذكاء) وحالات الرسوب والنجاح في امتحان الشهادة الاعدادية من ناحية ، والمجموع الكلي لدرجات هذا الامتحان من ناحية ثانية ، ودرجات مواد هذا الامتحان من ناحية ثالثة . وقد وجد هذا الباحث أن معامل الارتباط الثنائي بين نسب الذكاء وحالات النجاح والرسوب هو ٠.٥٤ ، ومعامل الارتباط بين نسب الذكاء والمجموع الكلي لامتحان الشهادة الاعدادية ٠.٨١٨ . أما معاملات ارتباط المقياس بدرجات المواد الدراسية فكانت في المتوسط ٠.٥٠ ولم يظهر التحصيل اللغوي (في اللغة العربية) معاملات ارتباط أعلى من التحصيل اللغوي (في الرياضيات) كما هو الحال في الدراسات السابقة ومنها بحث بوند (٥١) الشهير .

وأعد الباحث بعد ذلك مصفوفة معاملات ارتباط درجات الذكاء ودرجات المواد الدراسية المختلفة وخضعها للتحليل العاملي فوجد أن تشبع مقياس ستانفورد - بينيه بالعامل العام ٧٨٤ وهو أعلى التشبعات بهذا العامل ، إلا أن هذا ليس برهانا على صدق التكوين الفرض للمقياس لأن باقى المتغيرات كانت في الواقع درجات للتحصيل المدرسي ، ولم تتضمن المصفوفة أى مقياس آخر للذكاء .

وقد ظهرت على نسق الصورة المصرية للمقياس (طبعة ١٩٢٧) صورتان عربيتان أخريان أحدهما أردنية أعدها عبد الله زيد الكيلاني (١٢٦) والأخرى سعودية أعدتها آمال أحمد مختار صادق (٩) ، كما ظهرت صورة كويتية لطبعة ١٩٦٠ أعدها رجاء محمرد علام وكمال إبراهيم مرسى (١١٦)

كما قام عبد الله عبد العزيز المغيلي (٢٢) بتقنين طبعة ١٩٦٠ على عينة سعودية من ١٧٥ طلميذا بالمرحلة الابتدائية بالرياض تمتد أعمارهم بين ٦ ، ١٢ سنة ، وحسب له الصدق بمحك اختبار رسم الرجل والتي سبق تقنيه على البيئة السعودية وتراوحت معاملات الارتباط بين ٥٧ ر ، ٨٤ ر . للأعمار المختلفة ، أما الثبات فحسب بطريقة إعادة الاختبار - بفواصل زمنية طويلة - المبروعان ، وتراوح معامل الارتباط بين ٨٨ ر ، ٩١ ر ، وحسبت له معايير في صورة متوسطات عمرية تدل على العمر العقلي ، والسؤال الذي لم تجد له اجابة في هذا البحث هو : كيف استخدم الباحث متوسطات الأعمار للدلالة على العمر العقلي في مقياس ستانفورد - بينيه ، ومن المعلوم أن حساب العمر العقلي في هذا المقياس يقوم على أسس عرضناها بالتفصيل فيما سبق لا تعتمد على فكرة حساب المتوسط والانحراف المعياري ، ومن الطريف أن الباحث استخدم ضمن معايير تسمية الذكاء بالمعنى التقليدي ، (٢٢ : ١٢٩) في الوقت الذي افترض فيه هذا المعيار تماما من التراث السيكمي المعاصر .

وتكاد تكون الدراسة العربية الوحيدة - في حدود علمنا - التي قامت بتقنين طبعة ١٩٧٢ من هذا المقياس في دراسة جمال محمد على (١٢) ، وبدأ بالصورة المصرية القائمة والتي أعدها محمد عبد السلام أحمد ولويس كامل مليكة قطبها على عينة من تلاميذ مرحلة التعليم الأساسي في مصر (ن = ٢٠٧) تراوح أعمارهم بين ١٥ ، ٦ سنة ، ثم طبق على الأسئلة الطرق الاحصائية لتحليل الاحتمالات الاختيارية لكل سؤال ، وتحليل قائمة الكلمات وترتيب على ذلك تغيير وتعديل وحذف وإضافة بعض المفردات لكي تناسب طبيعة الثقافة والبيئة المصرية ، واستعان في كل ذلك بطبعة ١٩٧٢ من المقياس ، وانتهى به الأمر الى اعداد صورة مقترحة للمقياس طبقها على عينة أخرى من نفس المرحلة (ن = ١١٦) للتأكد من صلاحيتها للاستخدام وتلا ذلك تطبيقها على عينة ثالثة (ن = ١٠٠) لحساب الخصائص السيكميومية لها ، فحسب الثبات بطريقة كيودر - ريتشاردسون فبلغ معامل الثبات ٧٤ ر ، كما قدر الصدق بحساب معاملات الارتباط بين كل من الدرجات الخام ونسب الذكاء الانحرافية محكين هما تقديرات المعلمين ، ومواد التحصيل الدراسي والمجموع الكلي لها فتراوحت معاملات الارتباط بين ٢٠ ر ، ٥٥ ر وهي معاملات دالة احصائيا بالرغم من انخفاضها الظاهر . وقد قدر الصدق أيضا بطريقة

المجموعات المتضادة ، حيث قورن الأداء الاختبارى لعينة من المتخلفين عقليا والعاديين المساوين لهم فى العمر الزمنى ، ووجدت فروق دالة بينهما لصالح العاديين . وحسب الباحث للصورة المقترحة معاييرها فى صورة نسب ذكاء انحرافية .

أما طبعة ١٩٨٥ من المقياس فلا تكاد توجد دراسة عربية واحدة أجريت عليها .

مقاييس وكسلر للذكاء

ظهرت الطبعة الأولى من مقاييس وكسلر عام ١٩٢٩ وعرفت حينئذ باسم مقياس وكسلر - بليفو للذكاء Wechsler - Bellevue . وكان أحد الأغراض الهامة من أعداد هذه المقاييس توفير اختبار للذكاء يصلح للراشدين وقد أشار وكسلر فى ذلك الحين الى أن الاختبارات التى كانت متاحة حتى ذلك الوقت (ومنها بالطبع مقياس ستانفورد - بينيه) معدة فى جوهرها للأطفال وتلاميذ المدارس ، وكانت تستخدم مع المراهقين والراشدين عن طريق إضافة بعض الأسئلة الأكثر صعوبة من نفس النوع . ولذلك فإن من المشكلات الهامة فى هذه الاختبارات أن محتواها لا يثير اهتمام الراشدين (أى يعوزها الصدق الظاهرى) . وإذا لم يتوافر فى المقياس حدا أدنى من الصدق الظاهرى يصعب تكوين علاقة طيبة بين الفاحص والمفحوص الراشد .

وبالإضافة الى ذلك فإن المبالغة فى تأكيد عامل السرعة فى معظم الاختبارات الفردية يؤدى الى تعطيل أداء الراشدين . وكذلك فإن الاهتمام بالتناول الروتينى للنواحي اللفظية تحظى بالاهتمام الأكبر فى هذه المقاييس . ثم انه يستحيل استخدام معيار العمر العقلى مع الراشدين كما أشرنا . وأخيرا فإن عينات تقنين الاختبارات الفردية (مثل مقياس ستانفورد - بينيه) لم تتضمن الا قليلا من الراشدين .

ولمواجهة هذه المشكلات أعد وكسلر مقاييسه للذكاء ، ومن أسسها مقياسه لذكاء الراشدين . الذى يسمى فى الوقت الحاضر Wechsler Adult

أو Intelligence Scale WAIS ، والذي ظهر فى عام ١٩٥٥ ليحل محل مقياس وكسلر - بلفيو القديم . وقد ظهرت من المقياس طبعة جديدة معدلة فى أواخر السبعينات باسم WAIS-R لقياس ذكاء الراشدين من سن ١٦ الى سن ٧٤ سنة . وبالمطبع يتشابه المقياس الجديد فى الشكل والمحتوى مع المقياس السابق . الا انه تغلب على بعض الصعوبات الفنية الخاصة بمدى تمثيل عينة التقنين . وثبات الاختبارات الفرعية التى يتكون منها المقياس ، ثم ظهرت منه عدة طبعات آخرها ما صدر عام ١٩٨١ باستخدام مقياس وكسلر لذكاء الراشدين المعدل WAIS-R .

وقد ظهر مقياس وكسلر - بلفيو الأصيل باللغة العربية فى عام ١٩٥٦ وأعدده لويس كامل مليكه ومحمد عماد الدين اسماعيل (٢٨) ومنذ ذلك الحين لم تطرأ عليه تعديلات تذكر .

وفى عام ١٩٤٦ نشر وكسلر صورة ثانية لمقياس بلفيو الأصيل ، قننت على ألف راشد من الذكور أعمارهم تمتد من ١٨ - ٤٠ سنة ، الا ان هذه الصورة سرعان ما فقدت هويتها لأنها صارت أساسا لبناء مقياس آخر لذكاء الأطفال الذى شاع باسم مقياس وكسلر لذكاء الأطفال (WISC) والذي ظهرت منه طبعة معدلة عام ١٩٧٤ باسم (WISC-R) لقياس ذكاء الأطفال من سن ٦ الى سن ١٦ سنة .

وقد نقل المقياس الأصيل الى اللغة العربية محمد عماد الدين اسماعيل ولويس كامل مليكه عام ١٩٦١ ولا تتوافر عنه بيانات سيكومترية فى اللغة العربية حتى الآن .

وسرعان ما ظهرت الحاجة الى قياس ذكاء الاطفال فى سن ما قبل السادسة فأصدر وكسلر فى عام ١٩٦٧ اختبارا الثالث والهام والذي عرف باسم مقياس وكسلر لذكاء أطفال ما قبل المدرسة والذي عرف باسم (WPPSI) ويقس ذكاء الأطفال بين ٤ ، ٦ ١/٣ سنة .

وتتشابه المقياس الثلاثة فى البنية العامة حيث يتضمن المقياس مجموعة من الاختبارات الفرعية تشق منها درجة لفظية Verbal

ومجموعة أخرى تشتق منها درجة أدائية performance ، وترتبط الدرجتان معا في درجة كلية Full Scale . وتشابه الاختبارات الفسرعية في المستويات العمرية المختلفة الا انها لا تتطابق فهي تختلف في المحتوى والطول ومستويات الصعوبة والوزن النسبي للسؤال . ويوجه التطابق الوحيد هو المعيار المستخدم ، ففي الاختبارات الفرعية تحول الفسرجات الخام الى نمرجات معيارية معدلة متوسطها ١٠ وانحرافها المعياري ٣ ، اما في الاختبار الكلي فان الدرجة المعيارية المعدلة متوسطها ١٠٠ وانحرافها المعياري ١٥ .

ويوضح الجدول رقم (١٢) الاختبارات الفرعية التي تتألف منها المقاييس الثلاثة ويمكن الاستفادة به في المقارنة .

جدول رقم (١٢) الاختبارات الفرعية في مقاييس وكسلر الثلاثة

WPPSI	WISC-R	WAIS-R
مقياس ما قبل المدرسة	مقياس الأطفال	مقياس الراشدين
المقياس اللفظي		
المعلومات العامة	المعلومات العامة	المعلومات العامة
الفهم العام	الفهم العام	الفهم العام
الاستدلال الحسابي	الاستدلال الحسابي	الاستدلال الحسابي
المتشابهات	المتشابهات	المتشابهات
المفردات	المفردات	المفردات
(الجمل)	(مدى الأرقام)	مدى الأرقام
المقياس الأدائي		
رسوم المكعبات	رسوم المكعبات	رسوم المكعبات
تكميل الصور	تكميل الصور	تكميل الصور
—	ترتيب الصور	ترتيب الصور
—	تجميع الأشياء	تجميع الأشياء
بيت الحيوان	الترميز	رموز الأرقام
المتاهات	(المتاهات)	—
الرسوم الهندسية	—	—

(١) الاختبارات بين القوسين هي اختبارات احتياطية .

ونعرض فيما يلي وصفا مفصلا لكل مقياس من المقاييس الثلاثة ، ونبدأ بمقياس الراشدين ثم مقياس الأطفال ، أما مقياس أطفال مرحلة ما قبل المدرسة فسوف نتناوله مع الفقرة الخاصة بمقاييس ذكاء الرضع وأطفال ما قبل المدرسة في قسم لاحق من هذا الفصل . .

وصف مقياس وكسلر لذكاء الراشدين :

يتكون مقياس WAIS-R من ١١ اختبارا فرعيا ، ٦ منها تصنف في مقياس اللفظي verbal scale ، ٥ منها تصنف في مقياس أدائي أو عملي performance . وفيما يلي وصف موجز لهذه الاختبارات الفرعية مع بيان عدد الأسئلة التي يتكون منها كل اختبار في الطبعتين العربية والأجنبية .

المقياس اللفظي : يتكون المقياس اللفظي من ٦ اختبارات هي :

١ - المعلومات العامة : information ، ويتكون من ٢٩ سؤالا في الطبعة الأجنبية (٢٦ سؤالا في الطبعة العربية منها سؤال تمهيدي) تشمل مدى واسعا من المعلومات العامة التي يفترض شيوعها في ثقافة الراشدين . وقد روعي تحرر الاختبار من المعارف التخصصية والأكاديمية .

٢ - الفهم العام : comprehension ، ويتكون من ١٤ سؤالا في الطبعة الأجنبية (١٠ أسئلة في الطبعة العربية) ، وفي كل سؤال يشرح للمفحوص ما يجب عمله في ظروف معينة . ويتطلب هذا الاختبار قدرة على الحكم العملي والفهم العام Common Sense . وهو يشبه في هذا أسئلة الفهم في مقياس ستانفورد - بينيه ، إلا أن محتواه أقرب إلى اهتمامات الراشدين ونشاطهم .

٣ - الاستدلال الحسابي arithmetic ويتكون من ١٤ مسألة حسابية في الطبعة الأجنبية (١٠ مسائل في الطبعة العربية) تتفق مع مستوى مادة الحساب في المرحلة الابتدائية . وتعرض كل مسألة شفويا ، ويحلها المفحوص شفويا كذلك دون استخدام الورقة والقلم ، ولكل مسألة زمن محدد .

المتشابهات Similarities: ، ويتكون من ١٢ سؤالاً في الطبعة الأجنبية (١٢ سؤالاً في الطبعة العربية) تطلب من المفحوص أن يحدد كيف يتشابه شيئان (البرتقالة والموزة مثلاً) .

٥ - مدى الأرقام (أو إعادة الأرقام) : digit span ، وفيه تعرض على المفحوص قوائم تتكون من ٢ أرقام إلى ٩ أرقام ، ويكون عليه أعادتها شفويا ، وفي القسم الثاني من الاختبار يكون على المفحوص إعادة الأرقام بالعكس .

٦ - المفردات : vocabulary ، وفيه تعرض على المفحوص شفويا وبصرياً في وقت واحد ٤٢ كلمة متزايدة الصعوبة ، ويطلب منه تحديد معنى كل كلمة .

المقياس الأدائي أو العملي : يتكون المقياس العملي من ٥ اختبارات هي

٧ - رموز الأرقام digit symbol وهو عبارة عن صورة من اختبار التعريض الشفري code substitution المألوف في مقاييس الذكاء غير اللغوية . ويتكون مفتاح الشفرة من ٩ رموز تتزاوج مع الأرقام التسعة . ويعطى هذا المفتاح للمفحوص ويكون عليه أن يكتب أكبر عدد من الرموز التي تدل على الأعداد الموجودة في ورقة الإجابة في زمن لا يتجاوز ١٥ دقيقة .

٨ - تكميل الصور : Picture completion ويتكون من ٢١ بطاقة في الطبعة الأجنبية (١٥ بطاقة في الطبعة العربية) كل منها يحتوي على صورة ناقصة ، ويكون على المفحوص أن يذكر الجزء الناقص من كل صورة .

٩ - رسوم المكعبات : block design وتشبه مواد هذا الاختبار ما يوجد في اختبار Kohs Block Design Test ويتكون من صندوق به ١٦ مكعباً . و ٩ بطاقات على كل منها رسم مختلف (اثنتان من هذه البطاقات للتدريب) . وقد لون أوجه المكعبات بالألوان الآتية في الأصل الإنجليزي : الأحمر فقط والأبيض فقط والأبيض والأحمر معاً ، أما في الطبعة العربية فهي ملونة بالأزرق فقط والأبيض فقط والأصفر فقط والأبيض والأحمر معاً والأزرق

والاصفر معا . ويطلب من المفحوص انتاج رسوم على النحو الذى تحدده البطاقات ذات مستويات متزايدة من التعقد تتطلب ما بين ٤ مكعبات ، ٩ مكعبات . ولكل بطاقة زمن محدد .

١٠ - ترتيب الصور : Picture arrangement ، ويتكون كل سؤال فيه من مجموعة من البطاقات ، عليها صور ، ويطلب من المفحوص ترتيبها بالتتابع الصحيح بحيث تعطى قصة مفهومة . ولكل سؤال وقت محدد (عدد الأسئلة فى الطبعة الأجنبية ٨ وفى الطبعة العربية ٦) .

١١ - تجميع الأشياء Object assembly ، ويتبع هذا الاختبار نفس النمط المستخدم فى اختبار Pinter-Paterson الادائى . ويتكون فى الطبعة العربية من نماذج من الخشب لثلاثة اشياء هى الصبى او المانيكان، والوجه او البروفيل ، واليد وقد قطع كل منها الى قطع مختلفة ، ويطلب من المفحوص فى كل منها جمع القطع بحيث تكون الشكل الكامل .

تصحيح المقياس :

تصحح اختبارات الاستدلال الحسابى ورسوم المكعبات وترتيب الصور وتجميع الاشياء تبعاً لسرعة الأداء ودقته على النحو الآتى :

١ - الاستدلال الحسابى : تعطى درجة واحدة لكل مسألة تحل حلاً صحيحاً فى الوقت المحدد لها ، وتعطى درجة واحدة اضافية للزمن فى المسألتين ٩ ، ١٠ اذا حلت كل منهما فى حدود ٤٠ ثانية (الزمن العادى لكل منهما ١٢٠ ثانية) ، وتعطى درجتان اضافيتان اذا حلت المسألة فى حدود ١٥ ثانية ، بحسب الزمن بعد انتهاء المفحوص من قراءة المسألة (النهاية العظمى ١٤) .

٢ - رسوم المكعبات : تقدر الرسوم حسب الدقة والزمن الأقصى ، وتعطى درجات اضافية فى حالة اتمام الرسم فى زمن اقل . وقد أعد جدول لاعطاء تقديرات حسب زمن الانجاز (النهاية العظمى ٤٢) .

٢ - ترتيب الصور : تقدر الدرجة في هذا الاختبار على اساس دقة الترتيب والزمن المستغرق فيه ، وأعدت جداول لهذا الغرض بحيث تصحح كل من المجموعات الثلاث الأولى على أساس الكل أو لا شيء (صواب أو خطأ) . أما المجموعات الثلاث الأخيرة فتعطي تقديرات جزئية على كل ترتيب لا يطابق الترتيب الصحيح ، ولكن يمكن قبوله ، وبالنسبة للمجموعتين الأخيرتين تعطى تقديرات إضافية على الترتيب الصحيح اذ تم في زمن معين ، وتحسب الدرجة النهائية في هذا الاختبار بجمع التقديرات الجزئية في المجموعات المختلفة (والنهاية العظمى ٢١) .

٤ - تجميع الأشياء : يصحح الشكل الأول (الصبي) حسب الدقة فقط ، أما الشكلان الآخران (الوجه واليد) فيصحح كل منهما حسب الدقة والزمن معا . وفي الحالة الأخيرة تكون الدرجة في مواد الاختبار هي المجموع الكلي لتقديرات الدقة والزمن جميعا . وتوجد بكراسة التعليمات بروتوكولات خاصة لتصحيح كل شكل حسب الدقة ، وجدول للتصحيح على أساس الزمن (النهاية العظمى ٢٦) .

أما اختبارات المعلومات العامة ، ومدى الأرقام (أو إعادة الأرقام) ، وتكميل الصور ، ورموز الأرقام فإن الدرجة الكلية فيها هي مجموع الاجابات الصحيحة حسب مفتاح التصحيح المبين في كراسة التعليمات ، والنهايات العظمى لهذه الاختبارات الأربع على التوالي هي : ٢٥ ، ١٧ ، ١٥ ، ٦٧ .

وقد قام لريس كامل مليكه (٢٥) بإعداد نماذج تصحيح خاصة باختبارات الفهم العام ، والمتشابهات ، والمفردات ، نشرها في كراسة منفصلة عام ١٩٦٠ ، وفي هذه النماذج تعطى لأسئلة اختباري الفهم العام والمتشابهات درجتان أو درجة واحدة أو صفر حسب درجة التعميم في الإجابة ونوعه ، وبهذا تكون النهاية العظمى لكل من هذين الاختبارين هي : ٢٠ ، ٢٤ على التوالي . أما اختبار المفردات فإن أسئلته تعطى في نماذج التصحيح درجة واحدة أو نصف درجة أو صفر (النهاية العظمى ٢٤) .

المعايير :

طبقاته المختلفة من العينات القلائل التي يتوفر فيها شرط التمثيل ، فقد كانت تتألف من ١٧٠٠ شخصا في المقياس الاصلى ، اما المقياس المعدل الذي صدر عام ١٩٨١ فقد اعتمد على عينة تقنين قوامها ١٨٨٠ متساوي فيها اعداد الرجال والنساء موزعين في مستويات عمرية سبع بين ١٦ سنة و ٧٤ سنة اختيروا بحيث تتوافر فيهم خصائص التعداد العام للسكان في الولايات المتحدة الذي اجري عام ١٩٧٠ من حيث المنطقة الجغرافية والبيئة الريفية والحضرية والاصل العنصرى (وخاصة البيض والزنوج) ، والمستوى المهني والمستوى التعليمى ، واقتصرت العينة على ، الراشدين الأسوياء ، ، واستبعدت حالات التخلف العقلى ، والمصابون باصابات المخ ، والذين يعانون من اضطرابات سلوكية وانفعالية او نقائص جسمية ، كما استبعدت الحالات التي لا تستطيع استخدام اللغة الانجليزية .

وقد حولت الدرجات الخام في الاختبارات الفرعية الى درجات معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٢ . وكانت هذه الدرجات الموزونة Scaled Scores مشتقة من جماعة مرجعية عددها ٥٠٠ حالة شملت جميع المفحوصين من سن ٢٠ سنة ، ٢٤ سنة في عينة التقنين . وهكذا أصبحت الدرجات الفرعية في صورة وحدات قابلة للمقارنة ، ثم حسبت درجات المقياس اللفظى ، والمقياس الادائى ، والمقياس الكلى بجمع الدرجات الموزونة في الاختبارات اللفظية الست ، والاختبارات الادائية الخمس ، والاختبارات الأحد عشر جميعا على التوالى . ثم حولت هذه الدرجات الثلاثة الى نسبة ذكاء انحرافية مترسطة ١٠٠ وانحرافها المعيارى ١٥ في كل مجموعة عمرية ، وهكذا يتحدد الوضع النسبى للمفحوص بمقارنته بالأشخاص من نفس مستواه العمرى .

وقد استخدمت نفس الاجراءات في تقنين الطبعة العربية حيث استخدمت في اعداد الدرجات الموزونة ثلاث فئات عمرية هي :

١ - من ٢٠ الى أقل من ٢٥ سنة وعدد أفرادها ١١٤ .

٢ - من ٢٥ الى أقل من ٣٠ سنة وعدد أفرادها ١١٢ .

٣ - من ٣٠ الى أقل من ٣٥ سنة وعدد أفرادها ١٠٢ .

وقد استخدمت هذه العينة البالغ عددها ٢٢٨ مفحوصا فى اعداد نماذج ومعايير التصحيح لاختبارات الفهم والمتشابهات والمفردات بالاضافة الى اعداد جدول الدرجات الموزونة وجداول لنسب الذكاء الانحرافية لكل فئة من فئات السن (٣٥) .

وفى تصحيح هذا القياس تعطى للمفحوص درجات ثلاث : نسبة الذكاء الكلية ونسبة الذكاء اللفظية ونسبة الذكاء العملية . ويعطى فى الطبعة الأجنبية درجة رابعة تدل على ما يسميه وكسلر « معامل الكفاءة » ، efficiency quotient أو (E.Q.) ، ويتحدد بتقدير نسبة ذكاء الفرد فى الفئة من ٢٠ الى ٢٤ (فى اختبار وكسلر بلفيو طبعة ١٩٢٩) وفى الفئة ٢٥ - ٢٩ (فى اختبار وكسلر لذكاء الراشدين طبعة ١٩٥٥) ، بصرف النظر عن العمر الحقيقى للمفحوص . ويعتمد هذا المعامل على أساس ان الفئة من ٢٠ - ٢٤ (فى تقنين ١٩٢٩) والفئة من ٢٥ - ٢٩ (فى تقنين ١٩٥٥) هما اللتان وصلتا الى أعلى مستويات الأداء فى المقياس . ومعنى ذلك ان معامل الكفاءة طريقة للحكم على أداء الفرد فى ضوء متوسط أداء الأفراد من المستوى العمرى الذى يصل فيه الأداء الاختبارى اقصاه .

ويذكر لويس كامل مليكه (٣٥ : ٦١) انه يمكن الحصول على « معامل الكفاءة بصورة مؤقتة عن طريق استخدام جداول نسب ذكاء للفئة من ٢٠ - ٢٤ من العينة المصرية انفة الذكر ، لأن هذه الفئة حققت بالفعل أعلى مستوى من الأداء فى المقياس .

تفسير المعايير :

اعد وكسلر عام ١٩٤٤ فئات الذكاء التى يمكن ان تستخدم فى أغراض التصنيف الكلينيكى . وكان أساس التصنيف هو حدود الخطأ المعيارى للمتوسط . فالشخص الذى تكون درجته أقل من المتوسط بمقدار يبلغ ثلاثة أمثال الخطأ المعيارى أو أكثر يعد من فئة ضعاف العقول . أما الذى يحصل على درجة تقع ما بين خطاين معياريين وثلاثة أخطاء معيارية فيعد من قبيل الحالات الهامشية ، وهكذا . ويبين الجدول رقم (١٢) أسس تصنيف مستويات الذكاء فى مقياس وكسلر .

جدول رقم (١٢) تصنيف مستويات الذكاء فى مقياس وكسلر

التصنيف	حدود الخطأ المعيارى	حدود نسبة الذكاء الانحرافية	النسبة المئوية فى الأصل الاحصائى
ضعاف العقول	من ٢ فأقل	٦٥ فأقل	٢ر٢
الحالات الهامشية	من ٢ - حتى ٢	من ٦٦ حتى ٧٦	٦ر٧
الأغبياء	من ١ - حتى ٢	من ٨٠ حتى ٩٠	١٦ر١
المتوسطون	من ١ - حتى ١ +	من ٩١ حتى ١١٠	٥٠ر٠
أعلى من المتوسط	من ١ + حتى ٢ +	من ١١١ حتى ١١٩	١٦ر١
الممتازون	من ٢ + حتى ٣ +	من ١٢٠ حتى ١٢٧	٦ر٧
الممتازون جدا	من ٣ + فأكثر	من ١٢٨ فأكثر	٢ر٢

ثبات المقياس :

حسبت معاملات الثبات للطبعة الأجنبية من مقياس وكسلر لذكاء الراشدين لمجموعات من فئات الأعمار ١٨ - ١٩ ، ٢٤ - ٢٤ ، ٣٤ - ٤٥ ، ٤٥ - ٥٤ اختيرت من عينة التقنين على أساس تمثيلها لدى العمر الذى تشمله هذه العينة . واستخدمت طريقة التجزئة النصفية (بعد تصحيح سبيرمان - براون) فى حساب ثبات كل اختبار فرعى فيما عدا اختبارى رموز الأرقام ومدى (اعادة) الأرقام حيث استخدمت معهما طريقة الصور المتكافئة لعدم صلاحية طريقة التجزئة النصفية لهما بسبب غلبة عامل السرعة الشديدة عليهما .

وبلاحظ ان معاملات الثبات للمقياس الكلى قد بلغت فى المجموعات الثلاثة ٩٧ر ، وبلغت فى المقياس اللفظى ٩٦ر ، وفى المقياس الادائى ٩٤ر . ومعنى هذا ان المقاييس الثلاثة على درجة عالية من الثبات فى ضوء الاتساق الداخلى .

اما فى الطبعة العربية فقد حسب فرج عبد القادر طه (٣٠ : ٧٨) ثبات المقياس بطريقة اعادة الاختبار وطريقة التجزئة النصفية على عينة من عمال الصناعة ، وفى الطريقة الاولى اعيد تطبيق المقياس على ٤٠ حالة من

العمال فتراوحت معاملات الثبات بين ٦٠ر ، ٩٠ر . بالنسبة لكل من نسب الذكاء والاختبارات الفرعية . أما طريقة التجزئة النصفية فقد حسبت لدرجات ٧٠ عاملا وتراوحت معاملات الثبات الناتجة عنها بين ٥٠ر ، ٩٠ر . لنسب الذكاء والاختبارات الفرعية . ولم يوضح الباحث كيف استخدم طريقة التجزئة النصفية لاختبارى رموز الأرقام واجادة الأرقام بالرغم من عدم صلاحية هذه الطريقة لهما كما اشارت نتائج بحوث التقنين التى أجسريت على المقياس (٤٩ : ٢٤٧) .

صدق المقياس :

يفضل وكسلر (٨٢) فى تناوله لصدق مقياسه ان يؤكد « صدق المحتوى » . ومن زاوية ان الاختبارات الأحد عشر المتضمنة فيه تتفق مع تعريف الذكاء ، وأن اختبارات مماثلة استخدمت فى مقاييس الذكاء السابقة، وثبت جداولها فى الأغراض الكلينيكية وهذا القول اقرب الى صدق التكوين الغرضى ولكن بدون شواهد امبريقية . وبالإضافة الى ذلك يذكر وكسلر (٨٢) وماتارازو (٦٨) بعض البيانات عن الصدق التلازمى للمقياس . ومن ذلك المقارنة بين المجموعات التعليمية والمهنية المختلفة ، بالإضافة الى حساب معاملات الارتباط مع تقديرات الأداء على العمل والتقديرات المدرسية ، وكانت أغلب الفروق بين المجموعات فى الاتجاه المتوقع ، ومن ذلك ان درجات العمال أعلى فى الاختبارات العملية ، وان درجات الأفراد فى المهن غير اليدوية أعلى فى الاختبارات اللفظية . أما عن معاملات الارتباط بالمحكات المهنية والمدرسية فقد أعطى المقياس اللفظى قيما أعلى من المقياس الكلى أما المقياس الأدائى فقد أعطى أدنى القيم . وفى كل الأحوال كانت معاملات ارتباط المقياس اللفظى أدنى من معاملات مقياس ستانفرد - بينيه . أما عن الصدق التنبؤى للمقياس فقد ثبت أن المقياس يتنبأ بدرجة مقبولة بالضعف العقلى كما تؤكد بحوث جيرتن وزملائه (٦١ ، ٦٢) .

وقد حسبت معاملات الارتباط بين مقياس وكسلر وغيره من الاختبارات والمقاييس المعدة لمقياس الذكاء ، وكانت النتائج كما يلى :

١ - معامل الارتباط بين المقياس ومقياس ستانفرد - بينيه ٠٨٠ وهو أعلى فى المقياس اللفظى منه فى المقياس الكلى والمقياس الأدائى .

٢ - معاملات الارتباط بين المقياس واختبارات الذكاء الجماعية تتراوح بين ٤٠ ر ، ٨٠ ر .

٣ - معامل الارتباط بين المقياس الأدائي واختبار متيسورتا للوحات الأشكال ٧٢ ر .

٤ - معامل الارتباط بين المقياس الأدائي واختبار زافن للمصفوفات المتتابعة ٧٠ ر .

٥ - معامل الارتباط بين المقياس الأدائي واختبار بينت للفهم الميكانيكى ٣٥ ر .

صدق التكوين الفرضى : أجريت عدة تحليلات عاملية للاختبارات الفرعية التى يتألف منها مقياس وكسلر . ومن أشهر هذه الدراسات بحوث كوهن (٥٥ ، ٥٦) على مصفوفات ارتباطية لمجموعات عمرية أربعة من عينة التقنيين هى : (من ١٨ - ١٩ ، ومن ٢٥ - ٢٤ ، ومن ٤٥ - ٤٤ ، ومن ٦٠ - ٧٥ +) . وقد وجد الباحث أن الاختبارات الفرعية جميعا تشترك فى عامل واحد يسهم بحوالى ٥٠٪ من التباين الكلى للبطارية . وبالإضافة الى ذلك استخراج ثلاثة عوامل طائفية هى :

١ - عامل الذهن اللفظى وتشبع باختبارات المفردات والمعلومات العامة والفهم العام والمتشابهات .

٢ - عامل التنظيم الادراكى وتشبع باختبارات رسوم المكعبات وتجميع الأشياء ، وهو عبارة عن مركب من عامل السرعة الادراكية والتصوير البصرى المكانى اللذين اكدتهما البحوث العاملية الأساسية وبخاصة بحوث ثرستون .

٣ - عامل الذاكرة وتشبع باختبارات الاستدلال الحسابى ومدى (اعادة) الأرقام ، ويتضمن كلا من الذاكرة الصماء المباشرة للمواد الجديدة واستدعاء المواد التى سبق تعلمها .

ومن الطريف أن نذكر أن نتائج هذا البحث لم يدعم الاجراء الذي يستخدم فى تصحيح هذا المقياس أى تقسيمه الى مقياس لفظى ومقياس عملى ادائى لكل منهما نسبة ذكاء منفصلة . ورغم ان استخدام نسب الذكاء الكلية يدعمه وجود عامل عام مشترك فى جميع الاختبارات ، فان عامل الفهم اللفظى لم يظهر الا فى ٤ اختبارات لفظية فقط من ٦ اختبارات ، أما عامل الذاكرة فقد وجد فى الاختبارين اللفظيين الباقيين بالاضافة الى اختبارات فرعية أخرى من المقياسين عند الراشدين الأكبر سنا . أما عامل التنظيم الإدراكى فكانت تشبعاته الدالة فى اختبارين فقط من الاختبارات، الأدائية الخمس ، وأما الاختبارات الأدائية الثلاث الباقية فهى فيما يبدو لها عواملها الخاصة التى لا تشترك مع الاختبارات الأخرى المتضمنة فى بطارية المقياس .

وقد قام محمد عماد الدين اسماعيل (٤٢) بتحليل عاملى لمقياس وكسلر (الطبعة العربية) توصل منه الى عامل عام واحد مشترك فى جميع اختبارات المقياس لفئة السن من ١٦ - ١٩ ، وكانت أكثر الاختبارات تشبعا به اختبار المعلومات العامة ثم اختبار الفهم العام ، وتراوحت تشبعات الاختبارات الفرعية به بين ٩٥ ر ، ٦٤ ر . ومن الطريف أن الباحث يطابق بين العامل العام والذى وجدته فى بطارية وكسلر والعامل العام عند سبيرمان ، أما بالنسبة لفئة السن من ٢٠ - ٢٤ فقد وجد عاملين : عامل عام مشترك فى جميع الاختبارات ، وعامل ثنائى القطب تشبع به تشبعات موجبة كاختبارات المعلومات والفهم والمتشابهات والمفردات ، وتشبع به تشبعات سالبة كاختبارات ترتيب الصور وتكميل الصور وتجميع الأشياء ويمكن اعتباره عاملا يميز بين الاختبارات اللفظية فى مقابل الاختبارات غير اللفظية .

وأخيرا نذكر أن البحوث العاملية العديدة التى قام بها سوندرز (٧٤ ، ٧٥ ، ٧٦ ، ٧٧) ، مستخدما مصفوفات ارتباط الأسئلة الفردية للاختبارات الفرعية للمقياس ، توصلت الى ١٠ عوامل على الأقل ، ولم يجد الباحث تطابقا كاملا بينها وبين الاختبارات الفرعية للمقياس ، بل تأكد أن بنية بعض الاختبارات الفرعية على درجة كبيرة من « التعقد العاملى » ، وأن بعض العوامل تتعدى أكثر من اختبار فرعى واحد .

ولم يظهر لهذا المقياس صدق يذكر فى البيئات العربية الأخرى بعد

ظهر الصورة المصرية للمقياس ، باستثناء ما قام به فرج عبد القادر طه وصلاح احمد مرحاب (٢١) باعدادهما الصورة المغربية للمقياس ، وقاما بتقنينها على عينة من ١٥٠ مفحوصا من الجنسين ومن الأعمار من ٢٠ - ٢٤ سنة ، وحسب الثبات عن طريق اعادة الاختبار للاختبارات الأحد عشر وتراوحت معاملات الثبات بين ٧٧ر ، ٩٦٨ر ، وقدر الصدق باستخدام مجموعتين متضادتين من ضعاف العقول والأذكاء وكانت الفروق بينهما دالة فى المقاييس الثلاثة ، وأعدت جداول الدرجات الموزونة ونسب الذكاء الثلاثة .

وصف مقياس وكسلر لذكاء الأطفال :

يتكون مقياس وكسلر لذكاء الأطفال Wechsler Intelligence Scale for Children أو WISC من الأسئلة السهلة التى كان يتألف منها ، مقياس وكسلر - بلفيو الأصى . ويقسم المقياس - كما هو الحال فى مقياس الراشدين - الى قسمين : أحدهما لفظى والآخر عملى ، والمجموع الكلى للاختبارات الفرعية ١٢ اختبارا منها اختباران احتياطيان أو اضافيان عند الضرورة ، على النحو الآتى :

القسم اللفظى	القسم العملى (الأدائى)
(١) المعلومات العامة	(٦) تكملة الصور
(٢) الفهم العام	(٧) ترتيب الصور
(٣) الاستدلال الحسابى	(٨) تصميم (رسم) المكعبات
(٤) التشابهات	(٩) تجميع الأشياء
(٥) المفردات	(١٠) الترميز
(إعادة الأرقام)	(المتاهات)

وقد اعتبر وكسلر الاختبارات الاحتياطية هى تلك التى ارتبطت أدنى الارتباطات ببقية المقياس ، وعلى ذلك فإن اختبار مدى (إعادة) الأرقام فى المقياس اللفظى واختبار المتاهات فى المقياس العملى من هذا النوع . ويتطابق اختبار الترميز مع اختبار رموز الأرقام فى مقياس الراشدين فيما عدا غلبة الأسئلة السهلة عليه . أما الاختبار الوحيد الذى ليس له نظير فى (التقرير النفسى)

مقياس الراشدين فهو اختبار المتاهات ، ويتكون من ٨ متاهات ورقية متزايدة الصعوبة ، وتصحح في ضوء الزمن وعدد الأخطاء . وفى رأى وكسلر أن اختبارى الترميز والمتاهات يمكن أن يحل أحدهما محل الآخر ، ويتطلب اختبار الترميز زمنا أقل من اختبار المتاهات ، وقد يفضل بعض الفاحصين لهذا السبب .

تصحيح المقياس وتفسير الدرجات :

ولا يختلف مقياس الأطفال فى اجراءات تصحيحه من مقياس الراشدين الا اختلافات طفيفة . وتتلخص هذه الاجراءات فى تحويل الدرجات الخام فى كل اختبار فرعى الى درجات معيارية معدلة فى اطار المجموعات العمرية التى ينتمى اليها المفحوص . وقد أعد وكسلر جداول لهذه الدرجات الموزونة لفئات عمرية مقدارها أربعة شهور ما بين سن ٥ سنوات حتى سن ١٥ سنة . وتشبه الدرجات الموزونة هنا درجات مقياس الراشدين فى أن متوسطها ١٠ وانحرافها المعياري ٣ ، ثم تجمع الدرجات الموزونة وتحول الى نسب ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعياري ١٥ سواء لاسياس اللفظي أو المقياس الأدائي أو المقياس الكلى . وكذلك اهتم وكسلر فى المقياس الاصلى باقتراح بعض الطرق الاحصائية لحساب الأعمار العقلية المقابلة لدرجات مقياس الأطفال . ورغم ان معيار العمر العقلي ليس من خصائص نسب الذكاء الانحرافية التى تتألف منها معايير المقياس الا انه كان فى ذلك يواجه مطلبيا عمليا ملحا نتيجة لمشروع مفهوم العمر العقلي بين خاصة الناس وعامتهم بالنسبة لميدان قياس الذكاء . وقد أعدت بالفعل جداول لتحويل الدرجات الخام فى المقياس الى أعمار عقلية . الا ان هذه الحاجة لم تعد ضرورية فى الوقت الحاضر ولذلك نجد فى المقياس المعدل (WISC-R) يستبعد هذه الضرورة ، وبالإضافة الى ذلك فان المقياس المعدل يتخفف كثيرا من الأسئلة التى لا تتناسب مع الأطفال وأحل محلها أسئلة مشتقة من خبرات الطفولة .

كما زيد طول بعض الاختبارات الفرعية لزيادة الثبات ، كما طرأت تحسينات على طرق تطبيق المقياس واجراءات تصحيحه ، وقد أدى ذلك الى اعطاء متوسطات نسب ذكاء تقترب كثيرا من القيم التى نحصل عليها من

مقياس ستانفورد - بينيه (معايير ١٩٧٢) ، وبصفه عامة لوحظ أن معايير المقياس المعدل اعلى من معايير المقياس الاصلى .

وتتألف عينة تقنين مقياس وكسلر المعدل للأطفال (WISC-R) من ١٠٠ طفل و ١٠٠ طفلة فى الأعمار من ٦ر٥ سنوات الى ١٦ر٥ سنة ، وكان المجموع الكلى للعينة ٢٢٠٠ مفحوصا صنفوا على اساس الاحصاء العام للسكان فى الولايات المتحدة الذى اجرى عام ١٩٧٠ وفى ضوء نفس المتغيرات الذى ذكرناها بالنسبة لمقياس الراشدين ، وكانت جميعا من الاطفال الأسوياء .

ثبات المقياس : حسبت معاملات الثبات بطريقة التجزئة النصفية لكل اختبار فرعى فى مقياس وكسلر للأطفال ولكل من المقياس اللفظى أو المقياس العملى ، والمقياس الكلى وقد حسب الثبات بالتجزئة النصفية لكل مجموعة عمرية من المجموعات الاحدى عشرة الا ان الثبات بطريقة اعادة الاختبار حسب لثلاث مجموعات فقط هى : (١) من ٦ ١/٢ - ٧ ١/٢ (ب) من ١٠ ١/٢ - ١١ ١/٢ (ح) من ١٤ ١/٢ - ١٥ ١/٢ وقد بلغ متوسط معاملات الثبات بالتجزئة النصفية ٩٤ر للمقياس اللفظى ، ٩٠ر للمقياس العملى ٩٦ر للمقياس الكلى بينما بلغت هذه المعاملات بطريقة اعادة الاختبار ٩٣ر ، ٩٠ر ٩٥ر على التوالي .

اما معاملات ثبات الاختبارات الفرعية فكانت فى الطبعة المعدلة (طبعة ١٩٧٤) اعلى منها فى الطبعات السابقة ، فتراوحت معاملات الثبات لمختلف الأعمار بالتجزئة النصفية بين ٧٠ر و ٨٦ر ، وبطريقة اعادة الاختبار بين ٦٥ر ، ٨٨ر .

صدق المقياس : لا تتضمن كراسات تعليمات مقياس وكسلر للأطفال بيانات عن صدقه ، ولو أن بعض الشواهد تدل على ذلك فجداول المعايير لكل اختبار فرعى تظهر تمايزا فى العمر دون أن تقدم البيانات للقارئ فى ضوء هذا المحك . وقد قام عدد من الباحثين بحساب الصدق التلازمى للمقياس باستخدام محك التحصيل المدرسى وتراوحت معاملات الارتباط بين ٥٠ر ، ٦٠ر وكما هو متوقع كانت معاملات المقياس اللفظى اعلى . كما حسبت معاملات ارتباط مقياس ستانفورد - بينيه قبلت ٧٣ر بالنسبة للمقياس الكلى ، وكانت معاملات المقياس اللفظى مرة اخرى اعلى .

وتقدم كراسة تعليمات المقياس المعدل معلومات اضافية عن معاملات الارتباط بين الاختبارات الفرعية بعضها وبعض من ناحية، والمقياس اللفظي والمقياس العملي والمقياس الكلي من ناحية أخرى . وقد حسبت هذه المعاملات منفصلة لعينة قوامها ٢٠٠ مفحوص في كل مجموعته عمرية من المجموعات الاحدى عشرة في عينة القنين . وقد بلغ متوسط معاملات ارتباط المقياس اللفظي والعملي ٠٦٧ . ومعنى هذا أن جزئى المقياس بينها قدر مناسب من الاشتراك الا أن انخفاض معامل ارتباطهما نسبيا يبرر الابقاء عليهما كمقياسين منفصلين .

أما بحوث التحليل العاملى التى أجريت على المقياس بهدف تحقيق صدق التكوين الفرعى فقد استخرجت عوامل تشبه ما وجد فى تحليل مقياس الراشدين وهى : العامل العام ، والفهم اللفظي ، والعامل المكانى الادراكى . وعامل الذاكرة . وتأكد أن عامل الفهم اللفظي يرتبط ارتباطات جوهرية بالاختبارات الفرعية الخمسة فى المقياس اللفظي ، وأن العامل المكانى الادراكى يرتبط ارتباطا دالا باختبارى رسم المكعبات وتجميع الأشياء وهما من الاختبارات الفرعية فى المقياس الادائى العملى .

تقويم المقياس : يمكن القول أن مقياس وكسلر للأطفال مقياس جيد لذكاء الأطفال ويتفق فى الخصائص مع كثير من الاختبارات الفردية . وكانت استازى ترى فى طبعة ١٩٦٨ من كتابها (٤٩) أن الأساس المنطقى لهذا المقياس ظاهر التناقض . لقد ظهر مقياس وكسلر بلفيو الاصلى عام ١٩٢٩ لمقابلة الحاجة الى اختبار للراشدين لا يوفرها مقياس ستانفورد - بينيه أو أى مقياس آخر يعتمد على مجرد (رفع مستوى) مقاييس الأطفال ، وعندما حقق وكسلر هذا الهدف سعى الى اعداد مقاييس للأطفال تعتمد فى جوهرها على (خفض مستوى مقاييس الراشدين) بل أنه سار بالتناقض الى نهايته حينما اعد مقياسا للأطفال ما قبل المدرسة . الا أن هذا النقد لم يعد له ما يبرره بعد أن ظهرت الطبعة المعدلة من المقياس عام ١٩٧٤ وتضمنت محتوى « طفولى » تماما - الا أن أهم ما يلفت النظر فى هذا المقياس انه لم يحظ بدراسات تستحق الذكر عنه باللغة العربية ولا تكاد تتوافر بيانات عن معالمة السيكومترية يمكن الاعتماد عليها .

المقياس البريطاني للقدرة :

تولى مسئولية اعداد هذا المقياس ثلاثة من علماء النفس البريطانيين هم اليوت C.D. Elliott ومرتضى D.J. Murray وبيرسون L.S. Pearson وظهر باسم المقياس البريطاني للقدرة British Ability Scale وقامت بنشره المؤسسة القومية للبحوث التربوية (NFER) .

ويتألف هذه المقياس من ٢٤ مهمة يجد فيها معظم المتخصصين فائدة كبرى فى التشخيص لمستوى واسع من الأعمار ابتداء من سن ٢½ سنة وحتى ١٧ سنة .

وتصنف المهام تبعا لمدى الأعمار ، فالمهام الخاصة بالمزاوجة الإدراكية perceptual matching تقتصر على المدى العمرى من ٢½ الى ٩ سنوات، بينما لا تقاس « السرعة فى تجهيز المعلومات » قبل سن ٨ سنوات . ومن أمثلة المهام التى تقيس القدرة الأخيرة أن يعرض على المفحوص خمسة اعداد يتألف كل منها من ثلاثة أرقام ويطلب منه أن يختار اكبرها . وتوجد مهام أخرى لقياس الذاكرة والاستدلال الصورى بمعناد عند جان بياجيه بالاضافة الى الاستدلال الاجتماعى بالاضافة الى المعرفة الأساسية بالقراءة والكتابة والسهولة فى التعامل مع الأعداد .

والغرض الرئيسى من هذا الاختبار هو رسم بروفيل المفحوص . ويمكن للفاحصين استخدام بعض الاختبارات الفرعية دون البعض الآخر . ويرى مؤلفو الاختبار أن بعض الاختبارات الفرعية يمكن الحصول على متوسط درجاتها وتقدير « نسبة ذكاء عامة » منها والبعض الآخر يعطى « نسبة ذكاء بصرية » والبعض الآخر يعطى « نسبة ذكاء لغوية » وهو فى هذا أشبه بمقياس وكسلر ، وفى نفس الوقت فيه بعض خصائص مقياس ستانفورد - بينيه ولعل من أهم السمات الهامة فى هذا المقياس ان بعض اختبارات الفرعية يمكن أن تعطى وأن مفرداتها يمكن أن تختار فى ضوء أداء المفحوص ذاته . ومن ذلك مثلا أن اختبار « تصميم المكعبات » فيه يتألف من ١٦ نمطا ، ويمكن للطفل الصغير أن يعطى الأنماط الأحد عشر الأولى بينما يبدأ المراهق من النمط السادس حتى النمط الأخير . ويقدر مستوى قدرة المفحوص من خلال

سجل تراكمى لنجاحه وفشله بطريقة راش Rasch . ويبدو لنا ان هذا المقياس الجديد لا يزال فى حاجة الى مزيد من البحث ، كما ندعو الباحثين المصريين الى محاولة تطويره للثقافة المصرية خاصة والعربية عامة .

بطارية كوفمان :

فى عام ١٩٨٣ أصدر كوفمان وزوجته اختبارا هاما باسم بطارية تقويم الأطفال لكوفمان Kaufman Assessment Battery for children بحيث تصلح للأطفال من سن ٢ ١/٢ الى ١٢ سنة وتتألف البطارية من ١٦ اختبارا ، الا انها مثل المقياس البريطانى للقدرة ، تخصص بعض الاختبارات الفرعية للأطفال الصغار والبعض الآخر للأطفال الكبار .

وتقوم الأسس النظرية للبطارية على الاتجاهات الحديثة فى مجال سيكولوجية الجهاز العصبى وعلم النفس المعرفى ، ولهذا يمكن الحصول منه على أربع درجات ، كلية ، رئيسية هى :

١ - درجة التجبيز التتابعى Sequential Processing ويمكن الحصول عليها من أداء المفحوص لمهام الذاكرة والتي تتطلب منه اعادة انتاج سلسلة من الأعداد بعد الاستماع اليها ، أو اصدار سلسلة من حركات اليد بعد رؤيتها أو لمس رسوم يسهل التعرف عليها بنفس الترتيب الذى يستخدمه الفاحص فى تسمية الاشياء المصورة فى هذه الرسوم .

٢ - درجة التجبيز المتانى simultaneous Processing ويمكن الحصول عليها من أداء مهام تشبه ما تتضمنه اختبارات مشابهة لترتيب الصور فى مقياس وكسلر ، واختبار المصفوفات ، واختبار الاغلاق (الجشطالتي) . ومن الاختبارات الجديدة المستحدثة فى هذه البطارية اختبار يسمى ، النافذة السحرية ، وفيه يحرك الفاحص صورة بحيث تمر فى نافذة ضيقة ويطلب من المفحوص أن يتعرف على الشيء المصور .

٣ - الدرجة المركبة للتجهيز العقلى وتمثل متوسط الدرجتين السابقتين . وتمثل الدرجة التى يحصل عليها المفحوص فى الاختبارات

الفرعية التي يتم تطبيقها عن طريق التمثيل الصامت ما يسمى « الدرجة غير اللفظية » .

٤ - درجة التحصيل ويمكن الحصول عليها من أداء المفحوص على المهام التي تعتمد مباشرة على الخبرة الناتجة عن الثقافة أو المدرسة ، ويشمل ذلك القراءة والحساب وتسميه الصور التي تدل على أشخاص أو أماكن معروفة (اختبار المعلومات) أو تدل على أشياء (اختبار المفردات) .

وينعكس في هذه البطارية التمييز الشهير الذي اقترحه ريموند كاتل بين الذكاء « السائل » والذكاء « المتبلور » وخاصة في درجتى التجهيز الثانى والتحصيل على التوالى . أما التجهيز التتابعى فيدل على الكفاءة فى الانتباه والتسميع (البروفة الذهنية) وهى عمليات مستقلة عن الاستدلال والمعلومات . ولا يركز كرفمان على الدرجات المنفصلة للاختبارات الفرعية الفردية (كما هو الحال فى المقياس البريطانى للقدرة) ، ولكنه يقترح تفسير أى اختلاف ملحوظ داخل كل فئة من الفئات العامة .

ويستخدم المقياس فى مهايير الاختبارات الفرعية للتجهيز درجة معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣ كما هو الحال فى الاختبارات الفرعية لوكسلر ، أما بالنسبة للدرجات الكلية من ناحية ودرجة الاختبارات الفرعية للتحصيل من ناحية أخرى فإن متوسطها ١٠٠ وانحرافها المعيارى ١٥ . ويمكن للدرجات أن تتحول أيضا الى مئينيات ، والتي حسبت على المستوى القومى من ناحية وعلى أساس المستويات الاقتصادية الاجتماعية من ناحية أخرى .

ويمكن للطفل أن يستجيب على معظم اختبارات هذه البطارية بالإشارة أو غيرها من الاستجابات غير اللفظية ، ولا توجد مهمة فى البطارية كلها يتطلب أكثر من كلمتين كاستجابة لفظية . ولهذا يقترح المؤلفان استخدام اختبار الفهم من مقياس وكسلر كمكمل لاختبارهما . كما يؤكد دليل البطارية على أن الاختبارات تقيس « النشاط العقلى الراهن » وليس القدرات الفطرية .

قياس ذكاء الرضع والأطفال في سن ما قبل المدرسة

توجد فئة من اختبارات الذكاء الفردية أعدت خصيصا لقياس ذكاء الرضع والأطفال في سن ما قبل المدرسة . والواقع أن قياس الذكاء في هذه المرحلة له خصائصه ومشكلاته ، ولعل أهمها تدريب الفاحص تدريباً جيداً على أساليب الملاحظة العلمية المباشرة . وبالطبع يمكن تطبيق بعض الاختبارات الجماعية على أطفال هذه المرحلة ، إلا أنه بوجه عام يمكن القول أن الاختبارات الجماعية لا تصلح إلا بعد وصول الطفل إلى سن المدرسة (٦ سنوات) . ويترب على ذلك أن معظم اختبارات الذكاء في هذا السن (أن لم يكن كلها) من النوع العملي الشفوي وتتناول في جوهرها النمو الحسي الحركي ، والخلاصة أن اختبار الأطفال في هذه المرحلة (وبخاصة بعدما يصل الطفل إلى مستوى من النمو اللغوي يسمح بالاتصال والتواصل إنما هو عملية تفاعل بين الأفراد inter personal من نوع رفيع . ونعرض فيما يلي بعض الاختبارات المستخدمة في هذه المرحلة ونبدأ بأشهرها وهو مقياس وكسلر لذكاء أطفال ما قبل المدرسة .

مقياس وكسلر لذكاء أطفال ما قبل المدرسة :

ظهر مقياس وكسلر لذكاء أطفال ما قبل المدرسة (WPPSI)

أو Wechsler Preschool and Primary Scale of Intelligence

عام ١٩٦٧ ، ويهدف إلى قياس ذكاء الأطفال من سن ٤ سنوات إلى ٦ ½ سنة ، ويتألف من ١١ اختباراً فرعياً يستخدم منها ١٠ اختبارات فقط في تحديد نسبة الذكاء ، ومنها نجد ٨ اختبارات عبارة عن خفض لمستوى ، الاختبارات المناظرة لها في مقياس الأطفال WISC أنف الذكر ، أما الاختبارات الثلاث الباقية فقد أعدت خصيصاً لهذا المقياس . وتصنف الاختبارات الفرعية في هذا المقياس - كما هو الحال في مقياس WISC, WAIS إلى قسمين : لفظي وعملي ، على النحو الآتي (وقد وضعنا علامة * أمام الاختبارات الجديدة) .

المقياس اللفظي المقياس الادائي (العملي)

المعلومات	* بيت الحيوان
المفردات	نكلمة الصور
الحساب	المتاهات
المتشابهات	* الرسم الهندسي
الفهم	رسم المكعبات
* الجمل (اختبار اضافي)	

ويوصى وكسلر باستخدام الاختبارات اللفظية والعملية بالتناوب لتوفير التنوع والاحتفاظ بميل الطفل وتعاونه . ويتطلب المقياس في تطبيقه ما بين ٥٠ دقيقة و ٧٥ دقيقة في جلسة واحدة أو جلستين اختباريتين .

وقد حل اختبار الجمل في هذا المقياس محل اختبار مدى (اعادة) الأرقام في مقياس WISC ، وفيه يكرر الطفل كل جملة مباشرة بعد عرضها عليه شفويا . وهذا الاختبار يمكن ان يستخدم كبديل لاي اختبار لفظي آخر ، أو يطبق كاختبار اضافي لتزويد الفاحص بمعلومات أكثر عن الطفل ، وفي كلتا الحالتين لا يستخدم في حساب نسبة الذكاء . أما اختبار بيت الحيوان فيشبه اختبار رموز الأرقام في مقياس WAIS واختبار الترميز في مقياس WISC . ويتضمن الاختبار مفتاحا يعطى للطفل على لوحة يتألف من صور لكلب ودجاجة وسمكة وقطة ، لكل منها بيت (أسطواني) ملون بلون مختلف موضوع تحت الصورة ، وعلى الطفل ان يضع الأسطوانة (البيت) في ذات اللون المناسب في فتحة تحت صورة الحيوان على اللوحة . وتحدد درجة الطفل في ضوء الزمن وعدد الأخطاء وعدد الاستجابات المخرقة . أما اختبار « الرسم الهندسي » فيتطلب من الطفل محاكاة ١٠ رسوم بسيطة بقلم ملون .

تقنين المقياس : تم تقنين هذا المقياس على عينة تتألف من ١٢٠٠ طفلا مصنفة الى ٦ فئات ، مدى كل منها نصف سنة في الأعمار من ٤ الى ٦½ سنة وكانت كل فئة تتكون من ١٠٠ طفل و ١٠٠ طفلة اختيرت على أساس بيانات التعداد العام للسكان في الولايات المتحدة الأمريكية عام ١٩٦٠ .

وفى تصحيح المقياس يتم تحويل للدرجات الخام الى درجات معيارية معدلة متوسطها ١٠ وانحرافها المعياري ٣ ، اما مجموع الدرجات الموزونة فى المقياس اللفظى والمقياس العلى والمقياس الكلى فيتحول الى نسبة ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعياري ١٥ . بالاضافة الى الأعمار العقلية بالمعنى التقليدى :

وقد حسب ثبات الاختبارات الفرعية - فيما عدا اختبار بيت الحيوان - بطريقة التجزئة النصفية . أما اختبار بيت الحيوان فقد حسب ثباته بطريقة اعادة الاختبار لاعتماده الكبير على السرعة فى الأداء . وتراوحت معاملات الثبات للاختبارات الفرعية والمقاييس الكلية بين ٨٠ ر ، ٩٢ .

أما من حيث صدق المقياس فلزال الميدان يحتاج لمزيد من البحث ، وقد يرجع ذلك الى حداثة المقياس . وقد حسبت علاقته بمقياس ستانفورد - بينيه فبلغ معامل الارتباط ٧٥ ر فى مجموعة ٩٨ طفلا تتراوح أعمارهم بين ٥ ، ٦ سنوات . كما هو الحال فى مقياس WISC فان معامل ارتباط المقياس اللفظى بمقياس ستانفورد - بينيه اعلى من ارتباط المقياس العلى به . وقد أجريت عدة تحليلات عاملية للمقياس سجلها ساتلر (٧٣) وفيها تأكد وجود عامل عام بالاضافة الى عامل لفظى وعامل عملى وكان العاملان الطائفيان اقل ظهورا فى الأعمار الدنيا . وعمرما فان اجراءات تقنين هذا المقياس وحساب ثباته على درجة عالية من الدقة الفنية - كما تقول انستازى (٤٩ : ٢٩١) - وهو فى ذلك يتفوق على كثير من اختبارات ذكاء أطفال ما قبل المدرسة التى ظهرت قبل ذلك ، ويعكس التطورات التى طرأت على حركة بناء الاختبارات فى السنوات الأخيرة ، ومع ذلك فان هذا المقياس لم يظهر بأية صورة فى أى بيئة عربية على عكس مقياسى وكسلر السابقين للراشدين والأطفال .

الجدول الارتقائية لجيزل :

من المحاولات الكثيرة التى بذلت لمقياس ذكاء الرضع وأطفال ما قبل المدرسة ما يسمى بالمقاييس الارتقائية developmental scales التى تعتمد على نتائج البحث فى علم النفس الارتقائى . ومن أشهر هذه المقاييس

جداول جيزل Gesell schedules التي ظهرت عام ١٩٤٧ وتفيد في تحديد مستوى النمو الذي يصل اليه الطفل (بالمعايير الأمريكية بالطبع) . وتعتمد في تقدير جوانب النمو اجراء تقنين يعتمد على الملاحظة المباشرة لاستجابات الطفل لمثيرات مقفلة ، وتقويم سلوكه ونموه في مواقف الحياة اليومية فالملاحظة جوهر هذه الأداة بالاضافة الى المعلومات التي يمكن الحصول عليها من مصادر أخرى كالأم . وتشمل هذه الجداول فترة طويلة نسبيا من النمو حيث تمتد من سن ٤ أسابيع حتى سن ٥ سنوات . وتؤكد نتائج البحث ان هذه الجداول تعطى معاملات ثبات بين الفاحصين المدربين يزيد على ٩٥٪ على الرغم من ان نظام الملاحظة والتصحيح منها اقل تقينا اذا قورن بغيرها من الاختبارات فتدريب الملاحظين يرفع الثبات (٤٩ : ٢٥٦) . وقد استخدمت هذه الجداول منذ ظهورها من جانب الاخصائيين النفسيين والاطباء العقليين سواء في البحث العلمي أو الممارسة العملية . وتشمل هذه الجداول في صورتها الأخيرة خمسة حالات رئيسية هي : الحركات الغليظة ، الحركات الدقيقة ، اللغة ، السلوك التوافقي ، والسلوك الشخصي الاجتماعي والتي يزود الفاحص او الملاحظ مع كل منها بوصف لفظي ونصل للسلوك المعتاد في المراحل العمرية المختلفة بالاضافة الى رسوم توضيحية لهذا السلوك .

خفض مستوى مقياس ستانفورد - بينيه :

ادى هذا الاهتمام بذكاء الرضع واطفال ما قبل المدرسة الى ادخال تعديلات جوهرية على مقاييس بينيه بحيث يمكن خفض مستواها ، وكانت أولى هذه المحاولات ما قام به كولمان Kuhlman عام ١٩٢٢ حينما عدل مقاييس بينيه الأصلية بحيث تصل الى سن ثلاثة شهور ، وكان اختبارها من المحاولات المبكرة لمقياس ذكاء الرضيع ثم عدل هذا المقياس عام ١٩٢٩ وظهر باسم مقياس كولمان - بينيه بحيث يشمل معظم فترة الطفولة المبكرة .

من اهم المقاييس التي استخدمت أسلوب خفض مقياس ستانفورد - بينيه مقياس كاتل الذي يسمى مقياس كاتل لذكاء الرضيع Cattell Infant Intelligence Scale الذي يعتبره كثير من علماء النفس ملائما لمقياس ذكاء الرضع .

وقد أعد هذا المقياس بفرض ، خفض مستوى ، تعديل عام ١٩٢٧ من مقياس ستانفورد - بينيه ، وفيه استخدم كاتل أسئلة ستانفورد - بينيه بالإضافة الى أسئلة أخرى أنشئت أو اقتبست من جداول جيزل وغيرها من اختبارات الرضع المتاحة ، وصنفت الأسئلة الى مستويات عمرية ، وحسب العمر العقلي ونسبة الذكاء على النحو الذى استخدم فى طبعة ١٩٢٧ من مقياس ستانفورد - بينيه . ويشمل هذا المقياس مدى من العمر يمتد بين شهرين الى سن ٣٠ شهرا ، وخلال العام الأول تتحدد مستويات العمر فى فترات طول كل منها شهر واحد ، وفى العام الثانى تتحدد بفترات طولها شهران ، أما فى النصف الأول من العام الثالث فإن طول الفترات ٢ شهر ٠ فإذا اجتاز الطفل جميع الأسئلة المخصصة لهذه الأشهر الثلاثين الأولى من الحياة يستمر الفاحص معه بتطبيق مقياس ستانفورد - بينيه مبدئيا بمستوى ٣ سنوات .

وقد قنن مقياس كاتل لذكاء الرضع على مجموعة من ٢٧٤ طفلا ، أعيد اختبار كثير منهم فى الأعمار ٢ ، ٦ ، ٩ ، ١٢ ، ١٨ ، ٢٤ ، ٣٠ ، ٣٦ شهرا .

ومن الاختبارات التى تنتمى الى نفس الفئة اختبارات ميريل - بالمر Marrell Palmer واختبارات العالم الانجليزى فالانتاين C.W. Valentine والجداول الارتقائية التى أعدها فى بريطانيا العالم الانجليزى جـ ريفيت Griffith وفى رأى فرنون (٨٢ : ٦٣) أن جداول جريفيت هذه هى أكثر اختبارات ذكاء الرضع شمولا وأدقها تقنيا . فهى تتضمن ٣ أسئلة لكل أسبوع من أسابيع العمر فى العام الأول من الحياة ، وسؤالين لكل أسبوع فى العام الثانى ، وبهذا يبلغ عدد أسئلة القياس ٢٦٠ . وتصنف الى ٥ مجموعات ، كل منها متكافئة مع الفئات الأخرى فى الدرجات بحيث يعطى المقياس تخطيطا نفسيا (بروفيلا) للنمو فى كل فئة ، بالإضافة الى درجة كلية تسمى النسبة العامة General Quotient .

ويتضمن الجدول رقم (١٤) هذه الفئات الخمسة ، مع أمثلة منها فى بداية المقياس ونهايته .

جدول رقم (١٤) فئات مقياس جربيليت وأمثلة من أسئلته

الفئة	أمثلة من سلوك الطفل بعد الولادة	أمثلة من أسئلة المقياس عند اقتراب الطفل من سن عامين
(أ) النشاط الحركي	الرفس ، رفع الرأس	النزول على الدرج دون مساعدة أحد .
(ب) النشاط الاجتماعي	الابتسام ، التعرف على الأم	توجيه أسئلة اليه عن أشياء موضوعة أمامه على منضدة .
(ج) السمع والكلام	الاجفال من الضوضاء	استخدام جمل تتألف من ٤ كلمات أو أكثر
(د) تأزر العين واليد	تتبع أضواء معينة	قذف كرة في سلة .
(هـ) الأداء العملي	امساك الأشياء	تجميع أجزاء لعبة .

مقاييس بيلى لنمو الرضع :

شهد عقد الستينات والسبعينات اهتماما جديدا بالاختبارات التي تصلح للرضع وأطفال ما قبل المدرسة ، ومن أهم العوامل التي ساعدت على ذلك التوسع السريع في البرامج التعليمية المعدة للأطفال المتخلفين عقليا من ناحية والنمو الهائل في برامج ما قبل المدرسة في مجال التربية التعويضية للأطفال المحرومين اجتماعيا وثقافيا وقد ظهرت اختبارات جديدة لمواجهة هذه المطالب الملحة .

ومن أشهر هذه الاختبارات مقاييس بيلى لنمو الرضع Bayley Scales of Infant Development ويشمل مواد من جداول جيزل ومن بعض الاختبارات الأخرى التي تصلح لهذه المرحلة . وتتضمن مقاييس بيلى ثلاث أدوات متكاملة لتقدير المستوى الارتقائي للطفل من سن شهرين إلى ٢½ سنة وهي :

المقياس العقلى والمقياس الحركى وسجل سلوك الرضع • ويشمل المقياس العقلى عمليات الادراك والذاكرة والتعلم وحل المشكلات والتلفظ ، وبدايات التواصل اللفظى ، والتفكير المجرد فى صورة المبكرة • ويشمل المقياس الحركى القدرات الحركية الغليظة ومهارات المعالجة بالأيدي والأصابع ، ويتداخل هذا المقياس مع المقياس السابق حيث يلعب النمو الحركى دورا هاما فى تفاعل الطفل مع البيئة وبالتالي فى نموه العقلى أما سجل سلوك الرضيع فيستخدمه الفاحص أو الملاحظ لتقويم الجوانب المختلفة فى نمو الشخصية مثل السلوك الانفعالى والاجتماعى ، ومدى الانتباه ، والمتابعة • والتوجه نحو الهدف •

وقد أعدت لهذه المقاييس معايير مشتقة من أداء عينة ممثلة للمجتمع الأمريكى من ١٢٦٢ طفلا موزعين بالتساوى تقريبا حسب الأعمار من شهرين الى ٣٠ شهرا واستبعدت من العينة حالات أطفال المؤسسات والأطفال المتسررين والأطفال الذين يعيشون فى بيئات أسرية مزدوجة اللغة • ويعطى المقياس العقلى والمقياس الحركى مؤشرين منفصلين وتتحول الدرجات الخام فيهما الى درجات معيارية اعتدالية متوسطها ١٠٠ وانحرافها المعياري ١٦ (كما هو الحال فى مقياس ستانفورد - بينيه) - وقد أعدت هذه المؤشرات الارتقائية كمعايير داخل المجموعات العمرية مصنفة على فترات طولها نصف شهر خلال الأعمار من شهرين الى ٦ شهور ، وعلى فترات طولها شهر واحد للأعمار من ٦ شهور الى ٣٠ شهرا •

وحسبت معامل الثبات لهذه المقاييس باستخدام طريقة التجزئة النصفية لكل مجموعة عمرية ، وقد بلغ وسيط هذه المعاملات للمقياس العقلى ٨٨ ، وللمقياس الحركى ٨٤ ، ويعطى المقياس معامل اتفاق مرتفع بين الملاحظين والفاحصين • أما الصدق فلا تتوافر عنه أدلة كافية ، ولو أن بيلى يرى أن استخدام مقاييسه أفضل ما يكون لأغراض التشخيص وليس للتنبؤ •

مقاييس مكارثى لقدرات الأطفال :

من المقاييس الحديثة والهامة أيضا فى هذه المرحلة مقياس مكارثى لقدرات الأطفال **McCarthy Scales of Children Abilities** والذي يصلح للأطفال فى الأعمار من ٢½ الى ٨½ سنة • ويتألف من ١٨ اختبارا مصنفة

الى ستة مقاييس متداخلة هي : المقياس اللفظي والادراكي الاداني ، والكمي ،
والمعرفي العام ، ومقياس الذاكرة ، ومقياس الحركة .

ويقترَب المقياس المعرفي العام من فكرة المقياس التقليدي الكلي للنمو
العقلي ، وقيم الحصول عليه من مجموع درجات المقاييس الثلاثة الأولى
والتي تؤلف ١٥ اختبارا من ١٨ اختبارا تؤلف البطارية الكلية . وتتحول
الدرجات الخام في هذا المقياس الى درجات معيارية اعتدالية متوسطها ١٠٠
وانحرافها المعياري ١٦ حسب لكل مجموعة عمرية مداها ٢ شهر ، اما
درجات المقاييس المنفصلة فتتحول الى درجات معيارية اعتدالية متوسطها
٥٠ وانحرافها المعياري ١٠ لنفس المجموعات العمرية .

وقد بلغت عينة تقنين هذا المقياس ١٠٠ طفل تقريبا لكل مجموعة عمرية
من المجموعات العشر التي تمتد على فترات مداها نصف سنة للمستويات
العمرية من ٢ ١/٢ - ٥ ١/٢ سنة ، ومداها عام كامل للمستويات العمرية من
٥ ١/٢ - ٨ ١/٢ سنة ، وبلغ قوام العينة الكلية ١٠٣٢ طفلا سويا ، تمثل المجتمع
الأمريكي حسب تعداد ١٩٧٠ .

وحسب معامل الثبات بطريقة التجزئة النصفية فبلغ في المتوسط ٩٢
للمجموعات العمرية المختلفة في المؤشر المعرفي العام ، وبالنسبة للمقاييس
الخمسة الأخرى يتراوح معامل الثبات بين ٧٩ ر ، ٨٨ ر . واستخدمت أيضا
طريقة إعادة الاختبار ن (= ١٢٥ طفلا) فبلغ معامل الثبات للمقياس
المعرفي العام ٩٠ ر وتتراوح للمقاييس الخمسة الأخرى بين ٦٩ ر ، ٨٩ ر

أما الصدق فلا توافر عنه أدلة كافية باستثناء بعض البيانات عن
الصدق التنبؤي باستخدام محك التحصيل المدرسي في نهاية الصف الأول
بالمدرسة الابتدائية . وعموما فإن البطارية تعتمد في بنائها الأصلي على
مصادر متعددة : الخبرة الكلينيكية ونتائج البحث في علم نفس النمو ونتائج
بحوث التحليل العاطلي .

المقاييس المعتمدة على نموذج جان بياجيه :

ظهرت في السنوات الأخيرة مقاييس لمراحل العمر المختلفة معتمدة على

نتائج البحوث الرائدة التي قام بها عالم النفس السويسرى جان بياجيه لفترة تتجاوز نصف قرن . ومن السمات الهامة فى هذه المقاييس انها من النوع الرتبى ordinal ، وقد عرضنا بعض خصائص هذا النوع من المقاييس فى الفصلين الأول والسادس ، وبإيجاز نقول انها تفترض التتابع المنظم والموحد للنمو خلال المراحل المتتابعة ، وهى المراحل التى تمتد من مرحلة الرضاعة حتى مرحلة المراهقة وما بعدها ، والتي يسميها بياجيه المرحلة الحسية الحركية ، ومرحلة ما قبل العمليات ، ومرحلة العمليات العيانية ، ومرحلة العمليات الشكلية الصورية (راجع تفصيل هذه المراحل فى كتابنا المرجع رقم ١٨) . وتنقسم هذه الاختبارات أيضا بأنها مرجعة الى المحتوى وتنسب الى المستوى فى التفسير وليس الى المعيار ، لأنها تقدم وصفا كيفيا لما يستطيع الطفل أدائه فى كل مرحلة . وانهام التى تتألف منها هذه الاختبارات تهدف الى التعرف على مفاهيم معينة لدى الطفل ، أو ما يسميه بياجيه المخططات المعرفية Cognitive Schemata (١٨) وليس قياس سمات واسعة النطاق . وبالنسبة لتطبيق هذه الاختبارات فانه يتطلب الحصول على تفسير الطفل لحدث معين ملاحظ وأسباب هذا التفسير ويعتمد التصحيح فى جوهره على نوعية الاستجابة وكيفها وليس على عدد الاستجابات الصحيحة لعدد من الأسئلة ، كما هو الحال فى الاختبارات التقليدية . ولهذا نجد أن الاجابة الخاطئة التى تصدر عن الطفل فى الاختبارات المعتمدة على نموذج بياجيه تحتل أهمية خاصة كالاجابة الصحيحة فى فهم العمليات المعرفية لدى الطفل . وجوهر القياس هنا هو عملية حل المشكلة وليس نتائج هذا الحل .

وعلى الرغم من أن معظم الاختبارات من هذه الفئة لاتزال فى مرحلة البحث ، الا أنه توجد بعض الاختبارات تصلح للاستخدام العام نذكر منها :

١ - مقاييس النمو العقلى : من اعداد ليورفيـدو Laurendeau وبينارد Pinard خلال الفترة ١٩٦٢ - ١٩٧٠ ضمن مشروع ضخيم فى علم النفس النمو أجري فى جامعة مونتريال بكندا ويقيس بعض مفاهيم بياجيه الأساسية وخاصة السببية (العلية) والمكان . وتطبق الاختبارات بطريقة الاستخبار الشفوى Oral questionnaires - كما يسميها الباحثان - والتي تقع فى منزلة بين طريقة بياجيه غير المقتنة التى يسميها

• الطريقة الكلينيكية ، والطريقة المقننة الكاملة في الاختبارات التقليدية فجميع الأسئلة مقننة ، ولكن يمكن للفاحص أن يتخذ مسارات مختلفة بديلة من أجل مزيد من استطلاع واستكشاف عمليات التفكير لدى الطفل الفرد ، وذلك اعتمادا على الاستجابات المبدئية التي تصدر عنه .

• ويتم تصحيح البروتوكول الكامل لاستجابات الطفل لكل اختبار باعتباره وحدة وذلك في ضوء المستوى الارتقائي الذي يدل عليه كيف الاستجابات .
وقد أخضع الباحثان مقياسيهما للتحليل الاحصائي ، فاختارا عينة تقنين عددها ٢٥ طفلا ، ٢٥ طفلة في كل مستوى عمري من سن ٢ - ٥ سنوات ، وبلغ المجموع الكلي للعينة ٧٠٠ طفلا تمثل المجتمع الكندي المتحدث بالفرنسية في مونتريال .

٢ - المقاييس الرتبية للنمو النفسي : أعد هذه المقاييس أوزجيرس Uzgis وهنت عام ١٩٧٥ لأعمار أقل تصل إلى سن شهرين وتعتمد إلى سن سنتين فقط ، وهي مرحلة تشمل ما يسميه بياجيه المرحلة الحسية الحركية والتي تشتمل على ٦ مراحل (راجع وصفها بالتفصيل في المرجع رقم (١٨) .

وتشمل البطارية ٦ اختبارات تقيس ما يلي : ثبات الشيء ، تنمية وسائل للوصول إلى غايات بيئية مرغوبة ، المبادأة ، العملية الاجرائية ، العلاقات بين الأشياء في المكان ، نمو المخططات . وقد حصل الباحثان على معاملات اتفاق بين الملاحظين ، ومعامل ثبات باعادة الاختبار ، لكل مقياس والتي تراوحت بين ٨٠٢ر ، ٩٩١ر . كما حسب الباحثان الحد الأدنى لمعامل الرتبية Ordinality باستخدام طريقة جتمان المسماة Scologram فبلغ ٥٠ر . ألا ان استازي (٤٩ : ٢٧٤) يرى ان تطبيق هذه الطريقة على المقاييس المعتمدة على نموذج بياجيه موضع خلاف لا يتسع المقام لتناوله بالتفصيل .

٣ - مجموعة تقرير المفاهيم : وقد أعد هذه البطارية جولدشميد Goldschmid وبنتر Bentler عام ١٩٦٨ للأعمار من ٤ - ٧ سنوات ، وتقيس أحد المفاهيم الهامة في نموذج بياجيه وهو مفهوم ثبات (التكوين النفسي)

الخصائص Conservation (راجع ١٨) ، والذي يعد مؤشرا على تحول الطفل من مرحلة التفكير قبل العمليات الى التفكير باستخدام العمليات العيانية أو المحسوسة ، وهو التحول الذي يحدد بياجيه حدوثه فى سن ٧ - ٨ سنوات .

ولهذا المقياس ثلاث صور ، منها اثنتان متكافئتان فى قياس المكان الثنائى البعد والعدد والجوهر والكم المتصل والوزن والكم المنفصل ، وقد تأكد امبيرقيا انهما صورتان متكافئتان فى المتوسط والانحراف المعيارى ، ومعامل الارتباط بينهما ٠٩٥ . اما الصورة الثالثة فتقيس المساحة والطول ، ويتطلب تطبيق المقياس التزام الفاحص بتعليمات مكتوبة وتشمل عرض رسوم للمواد المستخدمة فى الاختبار وتوجيهات لمعالجتها وتناولها وتعليمات لفظية تقدم للمفحوصين .

وقد حسبت معايير فى صورة شينيات من افراد عينة تقنين تتألف من ٥٠٠ طفل وطفلة تمتد اعمارهم بين ٨ - ٨ سنوات تمثل مدينة لوس انجليس بولاية كاليفورنيا . وقد حسبت معاملات الثبات باستخدام ثبات المصححين ومعادلة كيورد - ريتشاردسون ، وطريقة اعادة الاختبار ، كما حسبت القياسية Smiability والرتبية للاختبار ، وبالإضافة الى وصف بنيته باستخدام التحليل العاملى . الا ان العينات كانت صغيرة ، ويحتاج الاختبار لمزيد من البحث . ولعل اهم النتائج التى يزودنا بها المقياس هى قابليته للاستخدام فى البحوث النفسية المقارنة او عبر الحضارية cross-cultural.

تعليق عام على المقاييس المعتمدة على نموذج بياجيه :

على الرغم من ان الاتجاه نحو بناء مقاييس نفسية تعتمد على نموذج بياجيه قد تأخر طويلا ، الا انه يمثل فى السنوات الأخيرة أحد معالم التطور التى يشهدها علم النفس المعاصر ، الا ان ذلك يجب الا يصرفنا عن بعض المشكلات المتضمنة فى هذه المقاييس ونلجأ أهمها ، ظهور بعض مظاهر « عدم الاتساق » ، decalage فى التتابع المتوقع للنمو ، ويتضمن هذا بعض الشك فى افتراضات الاتصال والانتظام والتتابع التى يطرحها نموذج بياجيه للنمو

العقلي . فكثيرا ما يلاحظ أن المرحلة تطابق أداء الطفل تختلف من مهمة لأخرى ، ليس فقط بالنسبة للعمليات المختلفة التي تتطلب للحل ، وإنما أيضا حين تستخدم نفس العملية في أنواع مختلفة من المحتوى .

الا أن ما نحب أن ننبه اليه أن هذه المقاييس ارتبطت ارتباطات مرتفعة مع اختبارات الذكاء التقليدية المقننة . ومرتبط بنفس الدرجة مع التحصيل المدرسي للاطفال في الصف الأول ، كما هو الحال بالنسبة لاختبارات الذكاء الفردية والجماعية . ومعنى ذلك أنه على الرغم من اختلاف المنهج فإن مهام بياجيه وأسئلة اختبارات الذكاء المقننة ، وأسئلة الاختبارات التحصيلية بينها جميعا قدر من التداخل والاشتراك ، وعلى الرغم من أن مهام بياجيه أصعب في التطبيق وتستغرق وقتا أطول ، إلا أن عائدها للتقويم النفسى كبير، فهي تقدم لنا وصفا كيفيا للنشاط العقلى للطفل لا توفره الأدوات الأخرى .

الاختبارات العقلية للذكاء

توجد طائفة من المقاييس العقلية تقيس الأداء العقلى ولا تتضمن أى اختبارات لفظية أو لغوية . وقد أشرنا الى أن مقاييس وكسلر تضمنت قسما عمليا أدائيا بالإضافة الى القسم اللفظى اللغوى ، وبالفعل فإن المقاييس العملية للذكاء استفادت من التطورات التي حدثت في ميدان القياس النفسى بعد ظهور مقياس وكسلر - بليفير في ميدان قياس الذكاء ، بالإضافة الى الجهود الى بذلت لقياس الذكاء ، العقلى ، . ولهذا يمكن القول انه ترجد فئة من اختبارات الذكاء العام تعتمد اعتمادا كليا على الأداء العقلى وحدد . وتتطلب نتيجة لذلك تطبيقا فرديا .

وقد يكون أول مقياس عملى للذكاء هو لوحة الأشكال التي ابتكرها سيجان Seguin عام ١٨٦٦ لتستخدم في قياس ذكاء ضعاف العقول، وشاعت هذه اللوحة بعد ذلك ، بل وأدخلت في كثير من اختبارات الذكاء العملية اللاحقة . وتتألف في الوقت الحاضر من ١٠ قطع خشبية توضع أمام الطفل بترتيب مقنن ، ثم يطلب منه وضع كل قطعة في موضعها من اللوحة بأسرع ما يمكن ، ويسمح له بثلاث محاولات لكل قطعة ، ودرجة المفحوص لكل قطعة هي أسرع زمن يضع فيه القطعة الخشبية من بين المحاولات الثلاثة .

وفى عام ١٩١٤ أعد نوكس سلسلة من الاختبارات العملية تطبق دون استخدام اللغة ، تضمنت لوحات اشكال متزايدة الصعوبة ، بالإضافة الى اختبارين أحدهما يسمى اختبار السفينة ، والثانى يسمى اختبار المكعبات ، ويتألف الأول من ١٠ قطع خشبية مستطيلة الشكل ترتب فى اطار خشبى لتكون فى النهاية صورة لسفينة فى البحر ، أما الثانى فهو اختبار للذاكرة المباشرة لسلسلة من الحركات تصدر عن الفاحص فى صورة فقر محدد تحديدا مسبقا بطريقة مقننة على كل مكعب من المكعبات الأربعة التى يتألف منها الاختبار، ثم يطلب من المفحوص أن يحاكيه .

وقد يكون مقياس بنتر - بيترسون Pinter-Paterson الذى ظهر عام ١٩١٧ أكثر الاختبارات أهمية من حيث التقنين . وهو يتألف من مجموعة من الاختبارات مشتقة من الاختبارات العملية السابقة ، وخاصة اختبارات سيجان وهيلى ونوكس وغيرها ، ويبلغ عددها الكلى ١٥ اختبارا (وتوجد صورة مختصرة للمقياس تتألف من ١٠ اختبارات فقط) .

وفى عام ١٩٢٢ ظهر اختبار كوهس لرسم المكعبات Kohs Block Design Test كمقياس يصاحب مقياس ستانفرد - بينيه فى التعرف على المتخلفين عقليا أو يحل محله . ويتألف من مجموعة من المكعبات المتماثلة الملونة سطوحها الست بالألوان الآتية : الأحمر ، الأزرق ، الأصفر ، الأبيض ، والأصفر والأبيض معا ، الأحمر والأبيض معا . وتعرض على المفحوص بطاقة - فى كل مرة - عليها رسم ملون ويطلب منه انتاجه عن طريق تجميع المكعبات المناسبة . ويختلف عدد المكعبات المطلوب لمختلف الرسوم تبعا لمستويات الصعوبة من ٥ مكعبات الى ١٦ مكعبا . وتعتمد الدرجة على زمن الاستجابة وصوابها .

ومن الاختبارات العملية للذكاء التى ظهرت فى الربع الأول من هذا القرن وذاعت شهرتها اختبارات متاهات بورتويس Porteus التى ظهرت لأول مرة عام ١٩١٤ ، وتتألف من سلسلة من المتاهات المرسومة على الورق مرتبة حسب الصعوبة ويمكن اعطاؤها دون تعليمات لفظية وذلك باستخدام المتاهات السهلة على سبيل التمثيل . وتصلح هذه الاختبارات للمستويات العمرية من ٣ سنوات حتى مستوى الرشد ، وفيها يطلب من المفحوص أن

يتتبع بالقلم الرصاص اقصر طريق من نقطة البداية الى نقطة النهاية في متاهة الورق ، دون رفع القلم مطلقا على الورقة . وحين يقع المفحوص في « خطأ » سواء خرج عن الطريق أو دخل في ممر مغلق يوقفه الفاحص عن الأداء ويعطيه محاولة أخرى على متاهة مماثلة ، فاذا أخطأ في المحاولة الثانية يعطى درجة « فشل » في هذا المستوى . ويسمح للمفحوص في المستويات العليا بأربع محاولات . وتعتمد الدرجة على الزمن المستغرق ، ويوضع في الاعتبار أثناء التصحيح رقم المحاولة التي اجتاز فيها المفحوص المتساهة بنجاح . ويؤكد بوربيوس أن اختباراتة تقيس قدرات الاستبصار والتخطيط التي تعجز الاختبارات اللفظية عن قياسها .

ويمكن القول أن أكثر بطاريات اختبارات الذكاء العملية شيوعا في الوقت الحاضر بطاريتان : أولاهما ظهرت في الولايات المتحدة عام ١٩٢٠ وتسمى اختبارات آرثر العملية ، وثانيهما اختبارات الكسندر .

وتتألف مطارية آرثر Arthur Point Scale of Performance Tests من أفضل الاختبارات العملية التي ظهرت قبل ذلك ، وتتكون من ٩ اختبارات هي :

- ١ - مكعبات فوكس .
- ٢ - لوحة الأشكال لسيجان .
- ٣ - متاهات بورتيوس .
- ٤ - اختبار كوهس لرسم المكعبات .
- ٥ - لوحة الأشكال التي تتألف من مربع وصليب ينقسم كل منها الى ٤ أقسام توضع فيها أجزاؤها .
- ٦ - لوحة الأشكال المتشابهة : وهي أكثر صعوبة لأن القطع الخشبية فيها أقرب الى التشابه .
- ٧ - المانيكان ويتألف من شكل خشبي لرجل تجمع أجزاؤه : الذراعان والساقان والراس والجذع . والبروقيل ويتألف من قطع خشبية تجمع معا لتكون وجه انسان في صورة بروكيل .

٨ - تكملة الصور (سهل) .

٩ - تكملة الصور لويلي Healy وهو اصعب من الاختبار السابق لأن القطع الخشبية فيه متماثلة .

وفي عام ١٩٤٧ ظهرت طبعة معدلة من مقياس آرثر الأدائي تتألف من ٥ اختبارات فقط تشمل تعديلات على اختبارات مكعبات نوكس ولوحة الأشكال لسيجان ومتاهات بورتوس ، بالإضافة الى اختبار آخر لتكملة الصور مأخوذ من سلسلة هيلي - فرنالد ، واختبار جديد يشبه اختبار كوهس لرسم المكعبات في أنه يطلب من المفحوص محاكاة رسوم متزايدة الصعوبة معروضة أمامه على بطاقات ، ويختلف عنه في أن الرسوم مصنوعة من الورق المقوى وملونة بالألوان عديدة . ويسمى هذا الاختبار الجديد باختبار آرثر لرسوم الاستنسل .

وفي عام ١٩٢٦ ظهر في بريطانيا اختبار مماثل هو مقياس كولنز - دريفر Collins-Drever ويتألف من ١١ اختبارا عمليا ، يضم الاختبارات التسع أنفة الذكر في مقياس آرثر (مع اعتبار البروفيل اختبارا منفصلا ، فتصبح ١٠ اختبارات) . بالإضافة الى اختبار لصورة نجمة خشبية مقسمة الى أجزاء يقوم المفحوص بتركيبها ، (وتوجد صورة أخرى له تتألف من ٧ اختبارات وتسليح للأطفال الصغار) .

أما الاختبار الهام الثاني فهو الاختبار الأدائي لالكسندر Alexander وهو محصلة الدراسات العالمية التي قام بها العالم الانجليزي الكسندر ، ويتألف من ٢ اختبارات ، اثنان منها هما اختبارا مكعبات كوهس وبناء المكعبات اللذان اشرنا اليهما ، بعد أن عدل الكسندر اجراءات تصحيحهما بحيث تسمح باعطاء أكثر من درجة واحدة للنجاح في رسم النماذج المختلفة وللزمن المستغرق . أما الاختبار الثالث فهو اختبار جديد أطلق عليه الكسندر اسم اختبار الازاحة Passalong ، ويتألف من ٩ صناديق صغيرة كل منها يحتوى على قطعة خشبية واحدة لونها أحمر وعدة قطع لونها أزرق (هذه القطع قد تكون على شكل مربعات أو مستطيلات) . وعن طريق تحريك القطع بالانزلاق ، دائريا فننتقل القطعة الحمراء من أسفل الى أعلى (دون

رفع القطع من أماكنها مطلقاً) • وقد رتبت مشكلات الاختبار تبعاً لمستويات الصعوبة ، وتصحح حسب الزمن المستغرق حتى الوصول إلى الحل الصحيح •

وتحول درجات هذا المقياس إلى أعمار عقلية ، وقد أعدت له معايير للذكور والإناث (الانجليز) من سن ٧ سنوات إلى ١٩ سنة • وتسمى نسبة الذكاء الناتجة • نسبة القدرة العملية Practical Ability Ratio وتصلح في رأي الكسندر في انتقاء الطلاب للتعليم الفني •

وبالطبع فإن معظم الاختبارات التي تستخدم في قياس ذكاء المعوقين تتخذ في جهرها صورة الاختبارات العملية ، ومن ذلك اختبار هيسكي - نبراسكا Hiskey-Nebraska لقياس ذكاء الصم الذي يتألف من ١٢ اختباراً فرعياً هي :

- | | |
|-------------------------|------------------------|
| ١ - أنماط الخرز | ٧ - أنماط المكعبات |
| ٢ - ذاكرة الألوان | ٨ - إكمال الرسوم |
| ٣ - تعيين الصور | ٩ - الغاز المكعبات |
| ٤ - تداعي الصور | ١٠ - التشابهات المصورة |
| ٥ - ثني الورق (أنماط) | ١١ - الاستدلال المكاني |
| ٦ - مدى الانتباه البصري | ١٢ - ذاكرة الأرقام • |

أما بالنسبة للمكفوفين قد ظهرت صور معدلة من مقياس ستانفورد - بينيه ومقاييس وكسلر ثلاثم ظروف أعاقاتهم •

الفصل الثامن

الاختبارات الجماعية للذكاء

اهمية الاختبارات الجماعية للذكاء :

ينتمى مقياس ستانفورد - بينيه وركسلر الى فئة الاختبارات الفردية للذكاء ، وهي اصلح ما تكون للاغراض الكلينيكية ، كما اشرنا في الفصل السابق . أما في برامج التقويم النفسى التى تتطلب اختبار أعداد كبيرة من المفحوصين فى الميادين التربوية والصناعية والادارية والعسكرية تصبح هذه الاختبارات الفردية أكثر تكلفة فى الجهد والمال والوقت ولذلك ظهرت الاختبارات الجماعية لمواجهة مثل هذه الصعوبات وبخاصة فى الحرب العالمية الأولى حين أعد الاخصائيون النفسيون العسكريون فى الولايات المتحدة اختبارا ألفا (وهو اختبار لفظى) واختبار بيتا (وهو اختبار غير لفظى) لأغراض الاستخدام الحربى .

خصائص الاختبارات الجماعية :

لا تتطلب الاختبارات الجماعية فى معظم الحالات فاحصا مدربا تدريبيا خاصا ، كما أن الاجابات فيها تسجل على أوراق منفصلة للإجابة ، ولهذا لا تتطلب هذا النوع من العلاقات الانسانية المباشرة بين فاحص ومفحوص .

والاختبارات الجماعية ميزات هامة منها تيسير اجراءات التطبيق ، واتساع حجم العينات التى يمكن أن تطبق عليها هذه الاختبارات وحسن تمثيلها للاصل الاحصائى السكانى العام ، ويسر وسائل التصحيح الموضوعية وتوافر معايير أكثر دقة ، وشمول عينات التقنين التى تصل الى أكثر من ربع مليون فى بعض الاختبارات الجماعية فى حين أنها لم تتعد ٤٠٠٠ مفحوص فى أحسن حالات الاختبارات الفردية .

إلا أن لهذه الاختبارات مشكلاتها أيضا ، ومن ذلك ضعف العلاقة بين

الفاحص والمفحوص ، وصعوبة تأكد الفاحص من تعاون المفحوص ، وعدم إحاطته بدوافع المفحوص لأخذ الاختبار أو ميله وإهتمامه به . بل يصعب على الفاحص أن يتعرف على المشكلات الطارئة على موقف الاختبار وخاصة بالنسبة للأحوال الداخلية للمفحوص أثناء الاختبار الجماعي . وكذلك فإن بعض المفحوصين يواجهون صعوبات في موقف الاختبار الجماعي ، وبالتالي يتعطل أداؤهم عنه في موقف الاختبار الفردي . كما توجد أدلة على أن الأشخاص المضطربين انفعاليا يؤدون في مواقف الاختبار الفردي أفضل من الاختبار الجماعي . وأخيراً فإن الاختبارات الفردية لا تيسر للفاحص الحصول على ملاحظات إضافية عن سلوك المفحوص أو تحديد العوامل المسؤولة عن ضعف الأداء في بعض الأسئلة أو بعض الاختبارات الفرعية .

وتختلف الاختبارات الجماعية عن الاختبارات الفردية في الغالب وفي طريقة ترتيب الأسئلة . فأغلب الاختبارات الجماعية الحديثة تعتمد على أسئلة الاختبار من متعدد ، وقليل منها يستخدم أسئلة النهايات المفتوحة . كما تصنف الأسئلة المتشابهة فيها إلى اختبارات فرعية منفصلة ترحبها للتعليمات . وفي كل اختبار فرعي ترتب الأسئلة عادة تبعاً لمستويات الصعوبة .

وقد تنشأ بعض الصعوبات العملية من استخدام نظام الاختبارات الفرعية المنفصلة . ومن ذلك أن الفاحص الأقل خبرة والأقل رقة قد يقع في أخطاء ضبط زمن الاختبار وخاصة في الاختبارات ذات الزمن القصير . وحتى يمكن تحقيق قدر من التوازن بين استخدام الزمن المهد للاختبار وتنظيم الأسئلة على نحر يسمح لجميع المفحوصين بمحاولة الإجابة على جميع أنوعها بمستويات صعوبة متزايدة يلجأ بناء الاختبارات النفسية في الوقت الحاضر إلى النظام الحلزوني الشامل Spiral-Omnibus وفيه تعرض أولاً أسهل الأسئلة من كل نوع ثم يتلوها الأصعب من كل نوع غالباً الأصعب ، وهكذا على نسق حلزوني تصاعدي في مستوى الصعوبة . وحتى يمكن تجنب الحاجة إلى إعادة التعليمات بالنسبة لكل سؤال ولخفض عدد مرات التحول في التأهب الناتج عن التعليمات (١٨) يلجأ بعض مؤلفي الاختبارات إلى استخدام النظام الحلزوني الشامل بالنسبة لمجموعات من الأسئلة ، يتألف كل منها من ٥ - ١٠ أسئلة .

ويرى بعض نقاد الاختبارات الجماعية أن هذا النوع يفرض قيودا على استجابات المفحوص ، وقد وجه هذا النقد خاصة الى أسئلة الاختيار من متعدد وإلى بعض الأنواع الشائعة من الأسئلة مثل القياس التمثيلي والمتشابهات والتصنيف . بل إن بعض هذه الأسئلة - فى رأى هؤلاء النقاد - تعاقب المفحوص المبتكر الذى يدرك تضمينات غير عادية فى السؤال . ويرد انصار هذه الاختبارات على هذا النقد بأن هذه الحالات نادرة الحدوث وخاصة اذا كانت أسئلة الاختبار حالت تحليلا دقيقا وتوافرت عنها بيانات صدق كافية .

وفى رأى فريز آخر من النقاد أن الاختبارات الجماعية لا توفر للمفحوص فرصا لتحليل الأخطاء والبحث عن أسباب اختيار المفحوص لاجابة معينة ، وهذا النقد صحيح بالطبع . ولهذا السبب وغيره لابد من أن يطبق مع اختبار الجماعى اختبارا فرديا على المفحوص أو أن يحصل على معلومات اضافية عنه من مصادر أخرى وخاصة فى الحالات التى تتطلب اتخاذ قرارات هامة بالنسبة له ، وخاصة فى الحالات الهامشية التى يصعب اتخاذ قرار واضح بشأنها .

وتوجد مشكلة أخرى بالنسبة لهذا النوع من الاختبارات وهى عدم المرونة التى تحدث فى أن كل مفحوص يختبر فى جميع الأسئلة . وفى هذا استهلاك للوقت دون مبرر ، والأجدى أن يركز كل مفحوص على الأسئلة المناسبة لمستوى قدرته ، وإذا حدث ذلك فإن المفحوص يتجنب الملل الناجم عن العمل مع أسئلة شديدة السهولة ، أو الشعور بالاحباط والقلق الناجم عن محاولة الاجابة على أسئلة شديدة الصعوبة وتتجاوز مستوى قدرته . والاختبارات الفردية فى هذا الصدد لها ميزة واضحة . ولعلنا نذكر أنه فى مقياس ستانفورد - بينيه يتم انتقاء الأسئلة التى تقدم للمفحوص على أساس استجاباته السابقة ، وتعطى له الأسئلة التى تقع فى نطاق مستوى قدرته .

دور الحاسب الالكترونى فى بناء نظام اختبارى تكيفى :

بذلت فى السنوات الأخيرة محاولات للجمع بين مزايا كل من الاختبارات الفردية والجماعية ، وخاصة فى مجال المرونة بين أسئلة الاختبار وخصائص

استجابات المفحوص فرد . ويسمى هذا الأسلوب بأسماء مختلفة منها النظام الاختبارى التكيفى adaptive أو التتابعى sequential أو المتفرع branched أو التفصيل tailored أو التفريدى individulized أو المبرمج programmed أو الدينامى dynamic أو المشروط بالاستجابة response-contingent (راجع انستازى ٤٩ : ٢٠١ - ٢٠٤) .
وهذه الأساليب أسير فى التناول حين تستخدم الحاسب الالىكترونى فى تطبيق الاختبار .

وتتلخص طريقة بناء النظام التكيفى (أو غير ذلك من الصفات السابقة) للاختبار فى البدء بإعطاء اختبار قصير يسمى الاختبار الاستكشافى routing test تشمل أسئلته مدى واسعا من الصعوبة . وفى ضوء استجابة المفحوص فى هذا الاختبار يوجه الى اختبار من بين عدة اختبارات ذات مستويات صعوبة مختلفة ، عادة ما يكون أطول من الاختبار الاستكشافى الا أنه يتفق مع مستوى قدرة المفحوص وتسمى هذه الطريقة « طريقة المرحلتين » .

وتوجد طريقة أخرى لبناء هذا النظام الاختبارى تسمى « طريقة الهرم » ، وفيها يبدأ جميع المفحوصين بالإجابة على سؤال متوسط الصعوبة ، فإذا كانت اجابة المفحوص عليه صحيحة يوجه الى سؤال أكثر صعوبة ، أما اذا كانت اجابته خاطئة يوجه الى سؤال أسهل ، ويتكرر ذلك بعد كل استجابة على كل سؤال .

وقد أستخدمت هاتان الطريقتان فى بناء اختبارات من نوع الورقة والقلم المعتادة أو من النوع الذى يتطلب استخدام الحاسب الالىكترونى . أما الطرق الأكثر تعقيدا والتي لا تتطلب انعطافا ثا محددة مقدما من تتابع الأسئلة فلا يمكن اجراؤها الا بالحاسب الالىكترونى .

نظرية البنية :

تستخدم إجراءات المعالجة على الكمبيوتر فى تطبيق الاختبار نظرية الاستجابة على المفردة item response theory وهى التى تسمى نظرية

السمة الكامنة latent trait theory في بعض الأحيان . وتقوم هذه النظرية على خاصيتين أساسيتين لنحنى المفردة أو السؤال وهما :

١ - مقياس القدرة : ويقدّر بالعدد الكلى للاستجابات على الاختبار .
وغير ذلك من المعلومات التي تتوافر للباحث عن استجابات عينة معينة من
المفحوصين على الاختبار .

٢ - احتمال حدوث الاستجابة الصحيحة للقدرة معينة : وهو دالة
لوضع المفحوص في مقياس القدرة . ويحسب هذا الاحتمال من البيانات
التي تمثل نسب الذين يجتازون السؤال أو المفردة من الأفراد ذوي المستويات
المختلفة في القدرة .

ويتألف النموذج الكامل لنظرية السمة الكامنة من ثلاثة بارامترات
تشتق رياضياً من البيانات الامبريقية عن مفردات الاختبار وهي :

١ - تمييز المفردة : ويتحدد رياضياً بميل slope المنحنى الدال
على علاقة مقياس القدرة باحتمال حدوث الاستجابة الصحيحة للمفردة ،
وهما اللذان يؤلفان المحور الأفقى والمحور الرأسى على التوالى فى هذا
المنحنى . ويرتبط هذا البارامتر عكسياً بالمسافة المطلوبة فى متصل مقياس
القدرة حتى يزداد احتمال حدوث الاستجابة الصحيحة للسؤال . وكلما
ازدادت قيمة التمييز يزداد الميل انحداراً .

٢ - صعوبة المفردة : ويتطابق هذا البارامتر مع موضع المفردة فى
محور مقياس القدرة حين يكون احتمال الاستجابة الصحيحة ٠.٥ .

٣ - التخمين : وهذا البارامتر يستخدم فى أسئلة الاختيار من متعدد
والتي من أجلها صممت خصيصاً نظرية السمة الكامنة . ويمثل احتمال
حدوث الاستجابة الصحيحة ببعض الصدفة ويتحدد بالخط المقارب الأدنى
lower asymptote فى المنحنى الخاص بالمفردة .

ويعتمد الحاسب الالكترونى على الخصائص السابقة التي تتوافر فى
مجموعة كبيرة من المفردات . فبعد أن يستجيب المفحوص على سؤال له

خصائص معينة ينفق الكمبيوتر السؤال التالي على أساس الاستجابات السابقة التي صدرت عن المفحوص لخفض أثر أخطاء التقدير . ويمكن أيضا أن تستخدم هذه الطريقة للمواءمة بين نقطة البداية في الاختبار أو السؤال المبدى فيه ومستوى قدرة المفحوص . وفي هذه الحالة يعتمد الحاسب الإلكتروني على البيانات السابقة المتوافرة عن المفحوص مثل نتائجه في الاختبارات الأخرى أو في الامتحانات المدرسية أو المعلومات التي يقدمها المفحوص عن نفسه في بداية جلسة الاختبار .

وتعتمد درجة المفحوص في هذه الحالة ليس على الأسئلة التي يجيب عنها اجابة صحيحة ، وإنما على مستوى الصعوبة والخصائص السيكمترية الأخرى للأسئلة .

بطاريات اختبارات الذكاء المتعددة المستويات :

تظهر مقارناتنا السابقة بين الاختبارات الجماعية من ناحية الاختبارات الفردية والاختبارات المعتمدة على الحاسب الإلكتروني من ناحية أخرى أن الاختبارات الجماعية للذكاء تقدم لجميع المفحوصين نفس المجموعة من الأسئلة بصرف النظر عن استجاباتهم الفردية . ولهذا السبب نلاحظ أن أى اختبار جماعى يجب أن يشمل مدى محدود نسبيا من الصعوبة بحيث يلائم مستوى معين من العمر أو الصف الدراسى أو القدرة . وحتى يمكن توافر مقاييس للنمو العقلى لمدى واسع نسبيا قام الباحثون فى هذا الميدان ببناء بطاريات متعددة المستويات multilevel . وبهذه الطريقة يتم اختبار المفحوصين فى المستوى الملائم لهم ، ويمكن للمستويات الأخرى أن تستخدم لاعادة اختبار نفس المفحوصين فى مراحل تالية ، أو فى أغراض التقويم النفسى المثارن بين مختلف الأعمار أو الصفوف الدراسية . ويتوافر فى هذه البطاريات أرضية وسقف مناسبان للأفراد الذى يقعون على أطراف التوزيع سواء بالنسبة للعمر أو الصف الدراسى وذلك نتيجة التداخل بين البطاريات المتتابعة . الا أننا يجب أن ننبه هنا الى أن المواءمة بين صعوبة المفردات ومستوى قدرة المفحوصين ليست بمستوى الدقة الذى يتوافر فى الاختبارات الفردية أو الاختبارات المعتمدة على الكمبيوتر ، وإنما عادة ما تكون هذه المواءمة تقريبية وعادة ما تتم فى ضوء المعلومات العقلية التى

تتوافر لدى الفاحص عن المفحوص وخاصة عمره الزمني أو صفه الدراسي وليس في ضوء استجاباته على أسئلة الاختبار .

واقضل مجال لتطبيق الاختبارات المتعددة المستويات هو المدرسة حيث تزداد الحاجة الى المقارنة بين المفحوصين خلال سنوات عديدة ، ولهذا السبب فان معظم المستويات في هذه البطاريات تتحدد في ضوء الصفوف الدراسية . ويتوافر في مستويات هذه البطاريات درجة مناسبة من المتصل سواء من حيث المحتوى أو العمليات النفسية والعقلية المقيسة . كما أن الدرجات التي يحصل عليها المفحوص من مختلف المستويات تتخذ نفس الصورة . بالإضافة الى أن عينات التقنين في هذه البطاريات تكون في العادة متكافئة على نحو لا يتوافر في الاختبارات التي تقنن منفصلة . وأخيرا فان بعض هذه البطاريات قد يمتد على نطاق واسع ابتداء من مرحلة الحضانه ورياض الأطفال حتى مستوى القبول في الجامعة .

وتصاغ معايير معظم هذه البطاريات في صورة نسب ذكاء انحرافية أو ما يشبهها من درجات معيارية معدلة ، وبعضها يستخدم عدة أنواع من المعايير كالمئينيات والمعايير السبائية ومعايير الفرق الدراسية . وتعطى اختبارات هذه البطاريات أنواعا مختلفة من الدرجات ومن ذلك الدرجة الكلية في البطارية بالإضافة الى درجات منفصلة للاختبارات اللفظية والكمية ، أو الأقسام اللغوية وغير اللغوية في البطارية .

ويختلف أسماء هذه البطاريات . وترصد أنستازي (٤٩ : ٢٠٥) ستة أسماء شائعة هي الذكاء والقدرة العامة والقدرة العقلية والنضج العقلي والاستعداد الأكاديمي والقدرة المدرسية ، وهي عبارات أقرب الى الترادف في الاصطلاح السيكمومتري ، ولو أن مصطلح « ذكاء » يقل استخدامه أكثر من غيره في السنوات الأخيرة ربما لأسباب ترجع الى المفاهيم المختلفة للذكاء التي عرضها كاتب هذا الفصل في موضع سابق (٢٤) . وفي رأي أنستازي (٤٩ : ٢٠٦) أن بطاريات المستويات المتعددة تتألف من « عينة من المهارات العقلية العظمى التي تعد متطلبات أساسية للعمل المدرسي وغيره من الأنشطة الهامة في الحياة اليومية » ، وبالتالي فهي تقيس استعداد الفرد للتعلم المدرسي - مثلا - في كل مرحلة من مراحل العملية التعليمية .

١٤ سنوات ذكاء أطفال دور الحضانة ورياض الأطفال والصفوف الأولى
من مرحلة التعليم الأساسى :

تعد دور الحضانة ورياض الأطفال ابتداء من سن ٥ سنوات والصف
الأول الابتدائى ابتداء من سن ٦ سنوات أدنى مستوى عمرى يمكن اعداد
اختبارات ذكاء جماعية له ، اما قبل ذلك فلا تستخدم الا الاختبارات الفردية
مثل ستانفورد - بينيه . وفى سن ٥ سنوات يمكن تطبيق الاختبارات الجماعية
على مجموعات من الأطفال لا تزيد عن ١٠ - ١٥ طفلا ، وفى هذه الحالة
يجب على الفاحص أن ينتبه للمفحوصين كأفراد للتأكد من اتباعهم
التعليمات ، ويمكن اختبار عدد أكبر من الأطفال اذا توفر أكثر من فاحص .

وفى اختبار الأطفال فى هذا المستوى يجب أن تعطى التعليمات شفويا
مع التوضيح اللازم وحل التمرينات والأسئلة التمهيدية (بشرط وجود عدد
كاف منها) . وعادة ما تتألف كراسة الأسئلة من صور ورسوم ، ولا تتطلب
من الطفل أكثر من وضع علامة تحت الاجابة الصحيحة . وقليل منها يتطلب
بعض الاستجابات التأزرية مثل رسم خطوط توصل بين نقطتين . ومعنى
ذلك أن اختبارات هذا المستوى لا تتطلب من الطفل اجادة القراءة او الكتابة .
ولذلك يمكن أن تنتمى الى فئة « الاختبارات غير اللفظية non-verbal
والتي يجب الخلط بينها وبين « الاختبارات غير اللفظية non-language
التي لا تتطلب من المفحوص معرفة باللغة سواء اكانت مكتوبة او منطوقة .
والتي تستخدم مع الأجانب او الأميين او الصم . فالاختبارات غير اللفظية
تتطلب من الطفل معرفة باللغة المنطوقة ، ولذلك فانها تصنف محترى الاختبار
ولا تصنف سلوك المفحوص . فالاختبارات الفهم اللفظى يمكن أن تطبق فى
هذا السن باستخدام الصور : وتقاس مفردات الطفل او فهمه للجمل بهذه
الطريقة . وتفضل انستازى (٤٩ : ٢١٩) - معنا للخلط - أن تسمى هذه
الاختبارات : الاختبارات « الصورة pictorial او الاختبارات التى
لا تعتمد على القراءة non-reading بدلا من عبارة الاختبارات غير
اللفظية .

ومن امثلة الاختبارات التى تصلح لهذا المستوى اختبار القدرة العقلية
لأوتس ولينون Otis-Lennon Mental Ability Test (المستوى الأول)

الذى ظهرت طبيعته الأخيرة عام ١٩٧٩ - ١٩٨٠ ، وله صورتان (ج ، ك) لكل المستويات التى يصلح لها هذا الاختبار (وتمتد الى نهاية المرحلة الثانوية) ويتكون المستوى الأولى Primary level من مستويين فى الواقع : المستوى ١ (لأطفال الحضانة) والمستوى ٢ (للنصف الأول من الصف الأول الابتدائى) . وهذان المستويان متطابقان فى المحتوى ، ولا يختلفان الا فى طريقة تعبير الطفل عن استجابته ، ففي مستوى الحضانة يضع الطفل دائرة حول الاجابة الصحيحة ، أما فى المستوى الثانى فان الطفل يسود بالقلم الرصاص مكان الاجابة الصحيحة فى ورقة الاجابة .

وتعطى تعليمات الاختبار شفويا ، كما ان لكل سؤال زمنا محددا (حوالى ١٥ ثانية) ويتطلب الاختبار ككل من ٢٥ - ٣٠ دقيقة ، ويطبق على جلستين بينهما فترة راحة قصيرة ، فى كل جلسة يعطى قسم من قسمى الاختبار . ويتكون القسم الأول من ٢٣ سؤالا من أسئلة التصنيف وفيها يطلب من الطفل ان يضع علامة تحت الصورة (أو الشكل الهندسى) الذى لا ينتمى الى فئة من ٤ صور (أو رسوم هندسية) ، أما القسم الثانى فيتكون من ٢٢ سؤالا تقيس التصور الذهنى اللفظى Verbal Conceptualization ، والاستدلال الكمي quantitative reasoning والمعلومات العامة ، والقدرة على اتباع التعليمات .

ومن أمثلة التصور الذهنى اللفظى ان يطلب من الطفل وضع علامة تحت الصورة التى تدل على نار . وفى الاستدلال الكمي يطلب منه مثلا وضع علامة تحت عدد النقط الذى يساوى عدد الأقسام المعطاة لدائرة مرسومة . وفى المعلومات العامة يطلب منه وضع علامة تحت صورة الشيء الذى نستخدمه فى الحديث مع الآخرين (تليفون) . وفى اختبار اتباع المعلومات يطلب من الطفل مثلا وضع علامة تحت الصورة التى تدل على زجاجة داخل مربع وعليها علامة نجمة . ويحصل المفحوص فى هذا الاختبار بمستوياته المختلفة على مؤشر للقدرة المدرسية .

وقد قنن هذا المقياس فى أمريكا (كما قننت مقاييس المستويات الأعلى التى يقيسها نفس الاختبار) على عينة كبيرة ممثلة للمجتمع الأسمى فى أمريكا ، ووضعت معايير فى صورة نسب ذكاء انحرافية ، انحرافها المعيارى (التقويم النفسى)

١٦ . كما أن له معايير مئينية وتساعية تتحدد فى ضوء فئة العمر والصف الدراسى . وللاختبار معامل ثبات مرتفع باستخدام طريقة كيودر - ريتشاردسون وطريقة اعادة الاختبار ، كما ان له صدق تلازمى وصدق تنبؤى جيد باستخدام درجات الاختبارات التحصيلية كمدك .

ويمكن استخدام هذه الاختبارات فى قياس قدرات الاستدلال المجرد عند الأطفال من ذوى صعوبات القراءة او التعويق التربوى من الصفوف الدراسية الأعلى .

اختبارات ذكاء قلاميذ مرحلة التعليم الأساسى :

من أشهر الاختبارات الجماعية التى تصلح ابتداء من الصف الرابع بمرحلة التعليم الأساسى المستوى (ب) من بطارية لورج - ثورنديك المتعددة المستويات Lorge-Thorndike Multi-Leve Eattery والتى ظهرت عام ١٩٦٤ . ويصلح هذا المستوى ابتداء من الصف الرابع الابتدائى . ولهذا الاختبار صورتان يتضمن كل منهما ٥ اختبارات لفظية و ٣ اختبارات غير لفظية . وتشمل الاختبارات اللفظية : المفردات ، وتكملة الجمل ، والاستدلال الحسابى ، والتصنيف اللفظى والمتماثلات . أما الاختبارات غير اللفظية فتشمل : تصنيف الأشكال ، وسلاسل الأعداد ومتماثلات الأشكال . ويتطلب الاختبار حوالى ٢٥ دقيقة للقسم اللفظى . و ٢٧ دقيقة للقسم غير اللفظى . ورغم أن لكل اختبار فرعى زمنه المحدد ، إلا أن هذه الاختبارات توصف بأنها من اختبارات القوة .

ومن البطاريات المتعددة المستويات التى لها أهمية خاصة اختبار القدرات المعرفية Cognitive Abilities Test من اعداد لورج وثورنديك وتتألف هذه البطارية الكلية من مستويين أوليين (ابتداء من مرحلة الحضانة ورياض الأطفال وحتى الصف الثالث من مرحلة التعليم الأساسى فى مصر) ومن طبعة متعددة المستويات تشمل ابتداء من الصف الثالث من مرحلة التعليم الأساسى وحتى نهاية المرحلة الثانوية . ويبدأ المفحوصون الذين يعطون مستويات مختلفة وينتهون منها باستخدام مجموعات مختلفة من الأسئلة ، أو بلغة البطارية « وحدات قياس » modules .

وتقيس الطبعة المتعددة المستويات ثمانية مستويات يتألف كل منها من عشرة اختبارات مصنفة الى ثلاث بطاريات هي :

١ - البطارية اللفظية وتتكون من اختبارات المفردات ، وتكملة الجمل ، والتصنيف اللفظي ، والقياس التمثيلي اللفظي analogy .

٢ - البطارية الكمية وتتكون من اختبارات العلاقات الكمية وسلاسل الأعداد ، وبناء المعادلات .

٣ - البطارية غير اللفظية وتتكون من اختبارات تصنيف الأشكال والقياس التمثيلي للأشكال ، وتركيب الأشكال .

وقد قننت البطارية على عينة جيدة التمثيل للمجتمع الأمريكي في المستويات الصفية العشرة المتضمنة فيها . وتحول الدرجات الخام فيها الى مقياس موحد سي جميع المستويات . والمعايير المستخدمة في كل بطارية هي الدرجات المعيارية الاعتدالية متوسطها ١٠٠ وانحرافها المعياري ١٦ (أي نسبة ذكاء انحرافية) ، كما تستخدم المئينيات والتساعيات ، ويمكن الحصول على مؤشر موحد من البطاريات الثلاث . وحسبت معاملات الثبات بطريقة كيودر - ريتشاردسون وبلغت في المتوسط ٩٠ ، كما استخدمت طريقة الصور المتكافئة وتراوحت معاملات الارتباط بين ٧٠ ، ٨٠ ، ٨٠ ، اما معاملات الثبات بطريقة اعادة الاختبار فكانت اقل . وبصفة عامة فان البطارية اللفظية اكثر ثباتا من البطاريتين الأخريين .

ومن حيث الصدق فان معاملات الارتباط بين درجات البطارية عالية وكذلك معاملات الارتباط بين الاختبارات الفرعية . وتؤكد التحليلات العاملية للبطارية وجود عامل عام في البطاريات الثلاث قد يدل على القدرة على الاستدلال باستخدام المحتوى المجرد والرمزي . وحسب صدق الاختبار أيضا باستخدام محكات متعددة منها مقاييس المهارات الأساسية والاختبارات التحصيلية والكفاءة التعليمية ، وكانت معاملات ارتباط البطارية اللفظية كالمعتاد اعلى من البطاريتين الأخريين . وعموما فان الصدق التلازمي اعلى من صدقها التنبؤي .

ومن الاختبارات المصرية التى تقيس الذكاء العام فى هذه المرحلة ما يلى :

١ - اختبار الذكاء الابتدائى : وهو اختبار مؤسس على اختبار بالارد Ballard الذى وضع فى انجلترا لقياس ذكاء التلاميذ بين سن الثامنة والرابعة عشرة من العمر . وقد نقله الى العربية اسماعيل القباني وأدخل عليه تعديلات « قليلة » (٨ : ٨) كان من الواضح « أنها ضرورية لجعله ملائما للبيئة المصرية » على حد قوله . وطبق ابتداء من عام ١٩٢٨ على نحو ٣٠٠٠ تلميذا بالمدارس « الابتدائية والثانوية والأولية والورش الصناعية ومدارس المعلمين والمعلمات الأولية » ، وقام القباني بتحليل أسئلته حتى وصل فى عام ١٩٢٠ الى الصورة النهائية للاختبار التى تتألف من ٦٤ سؤالاً (بدلا من ١٠٠ سؤال فى الأصل الأجنبى) مقسمة الى قسمين : الأول يتكون من ٣١ سؤالاً ، والثانى من ٣٣ سؤالاً . وفى كل منهما تتدرج الأسئلة حسب الصعوبة . وأسئلة الاختبار غير متجانسة ولكل منها زمن خاص . وقد طبق الاختبار المعدل على عينة تقنين كبيرة بلغت أربعة آلاف تلميذ من أعمار مختلفة وحسبت متوسطات الأعمار لأفراد كل عمر من ٧ سنوات حتى ١٥ سنة بفئات مقدارها شهرين ، وأعد جدول للأعمار العقلية . وهكذا فإن هذا الاختبار يستخدم مفهوم العمر العقلى ، ونسبة الذكاء بمعناها التقليديين عند بينيه وترمان فى النصف الأول من هذا القرن . وعلى أساس نسب الذكاء أعد القباني جدولا لتحديد مستويات التفوق والتخلف العقلين على النحو الآتى :

العبرى ١٤٠ فأكثر .

الذكى جدا من ١٢٠ - ١٤٠ .

فوق المتوسط من ١١٠ - ١٢٠ .

متوسط الذكاء من ٩٠ - ١١٠ .

دون المتوسط من ٨٠ - ٩٠ .

الغبى جدا من ٧٠ - ٨٠ .

ضعيف العقل أقل من ٧٠ .

كما أعد جدول يدل على مستوى الذكاء بطريقة لا تحتاج الى حساب العمر

العقلى . وذلك بترتيب التلاميذ الذين اختبرهم من كل عمر على حسب درجاتهم فى الاختبار ، ثم قسمهم الى فئات متفاوتة فى المستوى العقلى هى ١ ، ب ، ج ، + ، ج ، - ، د ، هـ على الترتيب . والفئة (ج) تدل على المتوسط بالطبع ، ويبين فى الجدول حدود الدرجات لكل طبقة من طبقات الذكاء . ولما كان تلاميذ هذه الفئات من عمر واحد تقريبا فانه كلما كان المستوى العقلى للفئة اعلى كان ذكاء افرادها اكبر . وبالطبع لا تطابق « طبقات » الذكاء هذه المراتب التى تتحدد بنسب الذكاء وان كانت تقترب منها كما هو مبين بالجدول رقم (١٤) .

جدول رقم (١٤) يبين طبقات الذكاء ونسب الذكاء المقابلة لها
(٨ : ٤٣)

نسب الذكاء المقابلة لها		الطبقة
الى	من	
٧٨	٠٠	هـ
٨٨	٧٩	د
٩٥	٨٩	ج -
١٠٤	٩٦	ج
١١١	١٠٥	ج +
١٢١	١١٢	ب
فما فوق	١٢٢	١

وقد حسب ثبات هذا الاختبار بطريقة اعادة الاختبار على مجموعات من التلاميذ تمثل اعمار ٩ ، ١٠ ، ١١ ، ١٢ فتراوحت المعاملات بين ٨٦ ، ٩٠ . اما الصدق التلازمى للاختبار فقد تحدد بمعامل الارتباط بينه وبين اختبار جوادنف لرسم الرجل فوجد ان معاملات الارتباط تراوحت بين ٤٧ر ، ٧٢ر لعينات من اطفال اعمار ١٠ ، ١٠ ، ١١ . ويزيد مقدار هذا الارتباط كلما صغر عمر الأطفال . كما حسب معامل الارتباط بين درجات الاختبار ودرجات متاهة بورتويس فبلغ ٥١ر لاطفال الصف الابتدائى ، ٤٧ر

لتلاميذ من عمرى ٩ ، ١٠ سنوات من فرق مختلفة ، بالاضافة الى دراسة علاقته بالنجاح المدرسى وتقديرات المدرسين للذكاء ، وكانت العينات فى كل الأحوال صغيرة العدد .

والواقع أن هذا الاختبار من الاختبارات الجيدة حقاً ، وكان يمثل بداية موفقة لحركة القياس العقلى فى مصر . وقد ظهرت طبيعته الثانية عام ١٩٢٩ ، ولو أنه توبع متابعة جادة كما حدث لقياس بينيه الاصلى فى فرنسا ، أو ستانفورد - بينيه فى امريكا ، أو بالارد فى انجلترا لقدم لعلم النفس فى مصر خدمة جليلة . وفى رأينا أنه لو توافر عليه جهد جديد لاعادة تقنيته وتعديله بعد انقضاء ما يقرب من نصف قرن على تقنيته الأخير لكان فى ذلك اسهام كبير لحركة القياس العقلى فى وطننا .

٢ - اختبار الذكاء المصور : هذا الاختبار من اعداد أحمد زكى صالح ، وجرب على مستويات عمرية مختلفة وتأكد أنه يصلح لقياس الذكاء ابتداء من سن الثامنة و تلك أعدت معايير الاختيار (فى صورة مثنيات ونسب ذكاء تقليدية) من سن ٨ سنوات حتى ١٧ سنة . وهو اختبار متجانس الأسئلة بمعنى أنه يتألف من ٦٠ سؤالاً كل منها يسأل المفحوص أن يبحث عن الشكل المخالف فى كل مجموعة من الأشكال تتكون من ٥ اشكال .

واستخرجت معاملات الثبات بطريقة التجزئة النصفية وتحليل التباين وتراوحت بين ٧٥ ، ٨٥ .

أما عن صدق الاختبار فقد ارتبط بالنتائج العقلية العام (كما يقاس باختبار القدرات العقلية الأولية) بمقدار ٣٤ ، وتشبع فى عدد من البحوث العاملة التى أشرف عليها مؤلفه بالعامل العام تشبعات تتراوح بين ٤٨ ، ٦١ . والاختبار فى رأى كاتب هذا الفصل (٢٤) يقيس إحدى قدرات الذكاء وهى التصنيف ويمكن أن يعد مقياساً لهذه القدرة وحدها ولا يتجاوز ذلك الى الذكاء بمعناه الشامل ، بالاضافة الى أن محتواه يتأثر بالعوامل الثقافية . وبالرغم من هذه الحدود التى تقيد استخدام هذا الاختبار كمقياس للذكاء إلا أنه لا يزال يستخدم فى كثير من البحوث الحديثة على أنه كذلك ، ناهيك عن قصور بيانات تقنيته التى لم تتجدد منذ أوائل السبعينات بعد وفاة مؤلفه .

٢ - اختبار الذكاء غير اللفظي : هذا الاختبار من اعداد عطية محمود هنا ، عن الأصل الاجنبى للاختبار العقلى المتعدد الغير لغوى لترمان ومككال ولورج • Non Language Multi-Mental Test

ويشبه هذا الاختبار الاختبار السابق (اختبار الذكاء المصور) فى انه يتطلب فى جميع أسئلته - وعددها ٦٠ - ان يحدد الطفل الشكل المخالف من بين ٥ اشكال ، ومعنى ذلك انه متجانس التكوين وقيس نفس عملية التصنيف التى أشرنا اليها بالنسبة للاختبار السابق • وحسب ثباته بمعادلة كيودر - ريتشارد سون لفئات أعمار من ٧ر٥ سنة الى ١٤ر٥ سنة فتراوحت بين ٧٢ ر ، ٠٨٢ أما صدق الاختبار فقد تحدد بمعاملات ارتباطه باختبار الذكاء الثانوى (للقبائى) فبلغ ٦٥ ر ، وقتن الاختبار على عينة من ٨٤١ طفلا فى المرحلتين الابتدائية والاعدادية تتراوح أعمارهم بين ٦ سنوات ٩ شهور و ١٦ سنة ، ١١ شهرا •

وحسبت معايير الاختبار فى صورة أعمار عقلية يمكن منها حساب نسب الذكاء بالمعنى التقليدى • وفى حالة الأطفال الذين يبلغون سن ١٤ سنة أو أكثر يكون العمر الزمنى فى معادلة نسب الذكاء مقدارا ثابتا هو ١٤ سنة أو ١٦٨ شهرا ، وهو إجراء تجاوزه المنطق السيكومترى الحديث • وينطبق على هذا الاختبار ما قلناه عن الاختبار السابق من حيث ضيق نطاق القدرات التى يقيسها من الذكاء وعدم تحديث بيانات تقنيته •

٤ - اختبار الذكاء الاعدادى : يهدف هذا الاختبار الى قياس القدرة على الحكم والاستنتاج خلال ثلاثة انواع من المواقف : مواقف لفظية ومواقف عددية ومواقف تتناول الأشكال المرسومة ، (٢٠ : ٤) وهو من اعداد السيد محمد خيرى الذى يرى انه يقترب من مفهوم العامل العام عند سبيرمان ، رغم انه اختبار غير متجانس التكوين • وهو بذلك أقرب الى اختبارات الذكاء من نمط ستانفورد - بينيه •

وقد بدأ تقنين هذا الاختبار عام ١٩٥٨ واشتملت عينة التقنين على ٢٤١٢ تلميذا من مدارس القاهرة و ١٨٠٣ تلميذا من مدارس الوجه البحرى و ٥٨٨ تلميذا من مدارس الوجه القبلى (أى أن العينة الكلية بلغت ٥٨٠٣ تلميذا) وحسب الثبات بطريقة اعادة الاختبار (ن = ٥١٤) فبلغ ٩٢ ر ،

وبطريقة التجزئة النصفية فبلغ ٠.٩١ أما في تحديد الصدق فقد حسب معامل ارتباط الاختبار باختبار الذكاء الابتدائي (للقباني) فوصل الى ٠.٦٥ .
بالإضافة الى حساب معامل ارتباطه بتقديرات المدرسين للذكاء فبلغ ٠.٥٤ .
وحسبت معايير الاختبار في صورة أعمار عقلية . وقد استخدم تقنين الاختبار في عدد من رسائل الماجستير والدكتوراه الا انه في حاجة الى إعادة تقنين .

اختبارات ذكاء طلاب المرحلة الثانوية والجامعات :

من أشهر الاختبارات التي تقيس ذكاء طالب المرحلة الثانوية اختبار المستوى الثاني من اختبارات قدرات المرحلة الثانوية والجامعة (SCAT) أو School and College Ability Tests وله صورتان ١ ، ب .
ويهدف هذا الاختبار الى التنبؤ بالتحصيل الأكاديمي ، ولذلك يعطى للمفحوص فيه ثلاث درجات . درجة في الاختبارات اللفظية ، ودرجة في الاختبارات الكمية ، ودرجة كلية . وتعتمد الدرجة اللفظية على اختبارين هما : فهم الجمل ومعاني الكلمات . وتعتمد الدرجة العددية على اختبارين أيضا هما : العمليات العددية والمسائل العددية (حل المشكلات العددية) .
ويتفق هذا الاختبار مع النظرية الحديثة لقياس ما يسمى « القدرات المستثمرة » (٢٤ : ٢٤) ، ولذلك فهو يعتمد على المعرفة اللغوية والعمليات الحسابية المتعلمة في المراحل التعليمية السابقة . وهو في هذا لا يختلف عن كثير من اختبارات الذكاء الأخرى التي تعكس مقادير متفاوتة من آثار « التعلم ، ولا تقيس « الوسع الفطري » مستقلا عن الخبرات السابقة .

ويمكن لدرجات الاختبارات اللفظية والكمية والكلية أن تصاغ في ضوء مقياس مشترك يسمح بالمقارنة بينها من مستوى لآخر من المستويات التي يقيسها الاختبار . كما يمكن لهذه الدرجات أن تتحول الى مئينيات أو تساعيات مشتقة من عينة تقنين شاملة لكل مستوى فيه .

وقد حسب ثبات الاختبار بطريقة كيودر - ريتشاردسون وتراوحت معاملات الثبات بين ٠.٨٨ ، ٠.٩٦ للمقاييس الثلاثة . وقد أجرى على هذا الاختبار تحليل دقيق للمفردات للحصول على تمييز واضح بين الفئتين

اللفظية والكمية . وتؤكد النتائج تمايز الاختبار تمايزا متصاعدا مع تزايد العمر والمستوى التعليمي بما يتفق مع نظريات التنظيم العقلي . أما عن صدق الاختبار فقد اهتم مؤلفوه بصدقه التنبؤي واستخدموا لهذا الغرض محكات التحصيل المدرسي ، وكانت معاملات الارتباط مع الاختبارات المدرسية أعلى منها مع التقديرات ، وكانت في المستويات الدنيا أعلى منها في المستويات العليا .

وقد أعدت مجموعة من الاختبارات العقلية لتستخدم في أغراض انتقاء طلاب الجامعات وتصنيفهم وإرشادهم وتوجيههم ، ومن أشهرها اختبار الاستعداد التعليمي Scholastic Aptitude Test (SAT) من إعداد هيئة امتحانات القبول بالجامعات في الولايات المتحدة (CEEB) College Entrance Examination Board والتي نشرت لأول مرة عام ١٩٢٦ . وتظهر صور عديدة جديدة من هذا الاختبار كل عام . ويحصل المفحوص على درجات منفصلة للقسم اللفظي والقسم الرياضي في الاختبار . ويعطى هذا الاختبار تقديرا مبدئيا لاستعداد طالب المرحلة الثانوية للعمل الجامعي .

ومن هذه الاختبارات أيضا امتحان برنامج القياس الجامعي American College Testing Program أو (ACT) ويتضمن ٤ أقسام هي : استخدام اللغة . استخدام الرياضيات . قراءة المواد الاجتماعية ، قراءة العلوم الطبيعية . وقد أعد هذا الاختبار لندكويست Lindquist . وبمعكس وجهة نظره في أن الامتحان يجب أن يكون عينة من العمل الجامعي . وتداخل فيه اختبارات الاستعدادات التقليدية واختبارات التحصيل ، ويركز على المهارات العقلية الأساسية المطلوبة للاداء الناجح في الجامعات .

ويوجد اختبار ثالث هو الاختبار السيكولوجي لجامعة أوهايو ، أو :

(OSUPT) Ohio State University Psychological Test

وهو اختبار قوة محض ، يطبق دون زمن محدد ، ويستغرق في العادة بين ١ ساعة وساعتين . كما أنه في جوهره اختبار لفظي يتضمن الاضداد

والمتمثلات اللفظية وفهم القراءة والعلوم الطبيعية والرياضيات والعلوم الاجتماعية والأدب .

ومن الاختبارات الحديثة في الظهور اختبار التأهيل الجامعي College Qualification Test ويعطى درجات في القسم اللفظي والعددي والمعلومات العلمية ومعلومات العلوم الاجتماعية ، والمعلومات الكلية ، والدرجة الكلية في الاختبار . وكذلك برنامج الاختبارات للكليات الأمريكية American College Testing Program (ACT) والذي يشمل اختبارات الاختبار الأكاديمي ويهدف من أربعة أقسام : الاستخدام اللغوي ، الاستخدام الرياضي ، قراءة الدراسات الاجتماعية ، قراءة العلوم الطبيعية .

ومن الواضح ان باستثناء اختبار SAT (الذي تكملة الاختبارات التحصيلية) فان جميع هذه الاختبارات تمثل مزيجا من الاستعدادات العامة ومعرفة مادة دراسية من المجالات الأكاديمية الأساسية . وحين تستخدم الدرجات المنفصلة فان صدقها الفارق في التنبؤ بالتحصيل في الميادين المختلفة يصبح موضع شك . ومعنى ذلك ان الدرجة الكلية هي افضل منبؤ بالأداء في جميع المقررات الجامعية تقريبا . واذا كان لابد من استخدام الدرجات الجزئية فان الدرجات اللفظية هي افضل المنبئات الفردية ، ويجب ان ننبه الى ان هذه الاختبارات لا تستخدم كبداية لاختبارات نهاية المرحلة الثانوية (الثانوية العامة في بلادنا) في التنبؤ بالتحصيل الجامعي ، فالواقع ان الدراسات الحديثة (٤٩ : ٢٢٢) تؤكد ان درجات المرحلة الثانوية يمكن ان تنبأ بالتحصيل الجامعي بنفس كفاءة هذه الاختبارات ، ان لم يكن افضل . وحين ترتبط درجات اختبارات الذكاء هذه بدرجات نهاية المرحلة الثانوية (الثانوية العامة) فان التنبؤ بالأداء الجامعي يصبح اكثر دقة .

ومن الاختبارات التي أعدت في مصر لقياس ذكاء تلاميذ المرحلة الثانوية وطلاب الجامعات ما يأتي :

(١) اختبار الذكاء الثانوي : وقد أعد أسماعيل القباني ويؤلف من ٥٨ سؤالا تتناول سلاسل الاعداد وأنتاج الجمل وإدراك العلاقات بين

الكلمات وإدراك السفافات والاستدلال . ويصلح للتطبيق على تلاميذ المرحلتين الإعدادية والثانوية في المدى العمري من ١٢ سنة إلى ١٨ سنة وعلى الكبار ممن تزيد أعمارهم على ١٨ سنة . وقد أعدت معايير الاختبار في صورة طبقات للذكاء هي : المستوى (أ) وهو أعلى المستويات ، ثم المستوى (ب ، ج ، د) ثم المستوى (هـ) وهو أدنى المستويات ، كما أعد جدول يتضمن الدرجات التي تحدد كل طبقة من كل مستوى عمري . ولا تتوافر لدينا بيانات سيكومترية عن هذا الاختبار .

(٢) اختبار الذكاء العالي : من إعداد السيد محمد خيرى على أساس مفهوم سبيرمان للعامل العام على أنه إدراك العلاقات والمتعلقات ويتشابه مع اختبار الذكاء الإعدادى في أنه يقيس القدرة على الحكم والاستنتاج خلال ثلاثة أنواع من المواقف : لفظية وعددية ومواقف تتناول الأشكال المرسومة . وقد قنن على عينة عددها ٥٨٢٨ طالبا في المدارس الثانوية والمعاهد العليا والجامعات . واشتملت العينة على مستوى الثانوى بفرعيه العلمى والأدبى ، كما اشتملت عينة المستوى الجامعى على كليات نظرية وأخرى عملية .

وحسب ثبات الاختبار بطريقتى إعادة الاختبار والتجزئة النصفية فبلغ معامل الثبات ٨٤٥ر ، ٨٨١ر على التوالى . أما في تقدير الصدق فقد حسب معامل ارتباطه باختبار الذكاء الثانوى فبلغ ٦٩٤ر ، كما حسب ارتباطه بتقديرات المدرسين لذكاء تلاميذهم في التعليم الثانوى فبلغ ٥٢٢ر . وأخيرا حسب ارتباطه بدرجات أفراد العينة الذين في السنة الإعدادية أو الأولى من المستوى العالى أو الجامعى في نهاية العام فكان ٥١٨ر . وأعدت للاختبار معايير في صورة مئينيات . وهذا الاختبار حاله كحال سابقه اختبار الذكاء الإعدادى يحتاج الى إعادة تقنين .

(٣) اختبار الاستعداد العقلى للمرحلة الثانوية والجامعات : هذا الاختبار من إعداد رمزية الغريب وينقسم الى ٥ أقسام هي :

البيضة الذهنية . التفكير الرياضى ، إدراك العلاقات المكانية ، فهم الرموز اللغوية ، التفكير المنطقى . وهو من اختبارات القوة ، ويحتاج في المتوسط الى ساعة ونصف للاقتناء منه . وقد حسب الثبات بطريقة

التجزئة النصفية ($n = ٥٥٠$) فبلغ معامل الارتباط ٠٩٢ر٠ واستخدمت طريقتان لحساب الصدق التلازمي هما حساب معامل ارتباطه باختبار للقدرات العقلية الأولية فبلغ ٠٧٧ر٠ ، والمقارنة بين المجموعات المتضادة وكانت الفروق بين الممتازين عقليا ، والمتخلفين عقليا لها دلالة احصائية .

اختبارات طلاب الدراسات العليا :

يمكن القول ان الاختبارات التي تستخدم في قياس طلاب الدراسات العليا تمثل مزيجا من اختبارات الذكاء العام واختبارات التحصيل . ومن اشهر هذه الاختبارات امتحانات سجل الدراسات العليا (GRE) Graduate Record Examinations والتي تطبق على الطلاب المتقدمين للالتحاق بالدراسات العليا في أمريكا . وقد ظهرت هذه السلسلة من الاختبارات عام ١٩٢٦ بهدف ترجية الطلاب وقياس كفاءة التدريس وانتقاء افضل العناصر للحصول على المنح والبعثات والتعيين في وظيفة معيد .

وينكون الاختبار في احدى صوره من اختبار للاستعداد واختبار تحصيلي في ميدان تخصص الطالب من مستوى رفيع مثل الاحياء والسياسة والفلسفة وعلم النفس واللغة الأسبانية ، الخ . والاختبار في جوهره اختبار استعداد دراسي او ذكاء عام ملائم لطلاب الدراسات العليا في بداية التحاقهم بها .

وللاختبار صورة أخرى لا تتألف فقط من اختبارات استعدادات وتحصيل رفيع ، وانما تشمل كذلك ثلاثة اختبارات من نوع اختبارات المجال ، هي العلوم الاجتماعية والانسانيات والعلوم الطبيعية . وهذه الاختبارات تؤكد فهم المفاهيم الأساسية والقدرة على تطبيق المعلومات وليس مجرد استدعاء الحقائق التفصيلية .

وتحول درجات هذا الاختبار الى درجة معيارية معدلة متوسطها ٥٠٠ وانحرافها المعياري ١٠٠ . وحسب له الثبات بطريقة كيورد - ريتشاردسون وطريقة التجزئة النصفية ، وتتراوح معاملات الثبات بين ٠٨٤ر ، ٠٩٥ر اما الثبات بطريقة اعادة الاختبار فيتراوح بين ٠٨٥ر ، ٠٩٧ر بفواصل زمني مقدار

عام واحد ، وبين ٧١ ر ، ٩٢ ر بفواصل زمنية مقداره عامان . وحسب الصدق
المتنبؤى فى ضوء محكات معينة مثل المدرجات الأكاديمية فى الدراسة العليا ،
والنجاح فى مقابل الفصل فى الدراسة ، وتقديرات الاساتذة ، وأداء الطلاب
فى امتحانات التأهيل للدكتوراه .

اختبارات ذكاء الراشدين المتفوقين :

يمكن القول أن اختبارات طلاب الجامعة والخريجين وطالب الدراسات
العليا ثلاث مستويات الراشدين المتفوقين لأغراض التقويم والبحث فى الميدان
المبنى . ومن ذلك أن اختبار ميللر للقياس التمثيلى Miller Analogies Test
وهو يصلح لتقويم طلاب الدراسات العليا ، يمكن أن يستخدم فى هذا الغرض .
ويتألف من متعائلات مرتبة مصاغة فى محتوى أكاديمى وعلى درجة كبيرة من
الصعوبة ، مما يجعله ينتمى الى فئة اختبارات القوة . ومن هذه الاختبارات
أيضا اختبار اتقان المفاهيم Concept Mastery Test وقد ظهر نتيجة البحوث
الطولية الرائدة التى قام بها ترمان فى ميدان المتفوقين ، ويتألف من أسئلة
القياس التمثيلى والمتشابهات والأضداد التى تنتمى الى محتوى متنوع فى
مجالات العلوم الفيزيائية والبيولوجية والرياضيات والتاريخ والموسيقى
وغيرها .

اختبارات الذكاء المتحررة من اثر الثقافة :

تسمى هذه الاختبارات الى التحكم فى بعض الأبعاد التى تتألف منها
الثقافة Culture ، والتى تختلف فيها الثقافات - وعلى وجه الخصوص
بعد اللغة . فاذا كان المفحوصون يتحدثون لغات مختلفة يجب أن يتحرر
الاختبار من هذا العنصر سواء من جانب الفاحص أو المفحوص . أما حين
تختلف الأطر التعليمية اختلافات واسعة وخاصة فى المجتمعات التى تنتشر
فيها الأمية ، كأغلب الدول النامية - فإن الاختبارات يجب أن تتحرر من عنصر
القراءة - كما أشرنا الى ذلك من قبل . وفى حالة الاختبارات المتحررة من
اثر القراءة Non-Reading لا يستبعد اثر اللغة ، الشفوية ، بالطبع ، لأن
هذه الاختبارات تعد لجماعات تتحدث بنفس اللغة . ومن العوامل الأخرى
التي تختلف فيها الثقافات عامل السرعة speed ، وتفاوت فى ذلك
ابتداء من الإيقاع الشخصى personal tempo وحتى الدافعية للعمل بسرعة

وتقدير العمل السريع ، وفى هذا تختلف الثقافات الفرعية للمجتمع الواحد (وخاصة بين الريف والمدينة) . وأخيرا فإن الثقافات المختلفة (فرعية أو عامة) تختلف فى التعامل مع محتوى الاختبارات . وقد حاول بعض المؤلفين أعداد اختبارات غير لغوية وغير قرائية ، ولكنها تذف من صور لأشياء واقعية كالآثاث والألعاب والآلات (راجع فى ذلك اختبارات المقياس العملى برنيسلر ، واختبارات الذكاء المصور والذكاء غير اللفظى) وبالمطبع فإن هذا المحتوى « السيمانتى » للصور يعتمد على اضرار الخبرة السابقة .

ولهذه الأسباب وغيرها - ظهرت اختبارات الذكاء المتحررة من أثر الثقافة Culture-free أو الاختبارات التى تصلح لأغراض المقارنة بين الثقافات Cross cultural وكان أول اختبار ظهر فى الميدان اختبار بينا الحربى أثناء الحرب العالمية الأولى . وفيه تم التحكم فى العنصر اللغوى سواء من جانب الفاحص أو المفحوص . وكانت تعطى التعليمات بالإشارة والإيماءات والتعميل ، أما استجابات المفحوص فكانت تتمثل فى مجرد وضع علامات بسيطة (فيما عدا اختبار واحد كان يتطلب من المفحوص كتابة الأعداد) . وقد أعيد تقنين هذا الاختبار - للأغراض غير العسكرية - عام ١٩٤٦ وأصبح يتلف من ٦ اختبارات فرعية هى : المناهات ، وتعويض الرموز والأعداد ، والسخافات المصورة ، ولوحة الأشكال المصورة ، وتكميل الصور ، والسرعة الإدراكية . ويتطلب فى هذا التعديل بعض التعليمات اللفظية إلا أنه يعتمد اعتمادا كبيرا على الأسئلة التدريبية .

وقد ظهرت عدة اختبارات فى هذا الميدان منها اختبارات سبيرمان الحسية للذكاء التى أعدها باللغة العربية اسماعيل القبانى وعبد العزيز القوصى ، المقياس الأدائى العالمى لليتر Leiter International Performance وسوف نقتصر على ٢ اختبارات هامة هى : اختبارات كاتل المتحررة من أثر الثقافة . واختبار المصفوفات المتتابعة لرافن واختبار رسم الرجل لجود انف .

(١) اختبارات كاتل غير المتحيزة للثقافة : أعد هذه الاختبارات العالم الأمريكى المعاصر ريموند كاتل ، ولها ثلاثة مستويات : المقياس الأول للأعمار من ٤ سنوات الى ٨ سنوات والراشدين المتخلفين عقابا ،

والمقياس الثانى للاعمار من ٨ سنوات الى ١٢ سنة والراشدين العاديين ،
والمقياس الثالث من سن ١٢ سنة الى سن ١٦ سنة والراشدين المتفوقين .
ولكل مقياس من هذه المقاييس الثلاث صورتان متكافئتان . ويتطلب المقياس
الأول تطبيقاً فردياً أو جماعياً . ويتألف المقياس الأول من ٨ اختبارات ، أربعة
منها فقط تعد من نوع الاختبارات المتحررة ثقافياً . أما الاختبارات الأربعة
الأخرى فتتطلب كلا من الفهم اللفظى والمعلومات الثقافية الخاصة . أما
المقياسان الثانى والثالث فمتشابهان فى المحتوى ومختلفان فى مستوى
الصعوبة ، ويتألف كل منهما من أربعة اختبارات هى : سلاسل الأشكال ،
وتصنيف الأشكال ، ومصفوفات الأشكال ، والشروط . وتتوافق للمقياس الأول
فقط معايير نسب ذكاء تقليدية ، أما المقياسان الثانى والثالث فمعاييرهما من
نوع نسب الذكاء الانحرافية بانحراف معيارى ٢٤ أو ١٦ . وقد شاع التحويل
الآخر حتى يسمح للمقاييس بالمقارنة بغيرها .

وقد نقل المقياس الثانى الى اللغة العربية احمد عبد العزيز سلامة
وعبد السلام عبد الغفار ، ولا تتوافق عنه - فى حدود علمنا ببيانات تقنين ، أما
المقياس الثالث فتقام بأعداد دوزاد أبو حطب وآمال أحمد مختار صادق وقام
بتقنيته عدد من طلاب الدراسات العليا فى مصر فى بحوثهم للماجستير
والدكتوراه . كما استخدمت آمال صادق المقياسين الثانى والثالث فى
دراستين عاملتين للقدرة الموسيقية (٩) وأكدت النتائج تشبع المقياسين
بالمعامل العام فى عينتين مصريتين أحدهما من الأطفال والأخرى من المراهقين
والراشدين . ومقياس هذه الاختبارات للمعامل العام هو الغرض الأساسى
الذى تقوم عليه اختبارات كاتل متأثراً فى ذلك بنظرية أستاذه سبيرمان .

(٢) اختبارات المصفوفات المتتابعة لرافن : ظهر هذا الاختبار فى
انجلترا لمقياس العامل العام عند سبيرمان أيضاً ، وتتطلب فى جوهرها ادراك
العلاقات بين وحدات مجردة ، وهو فى رأى معظم علماء النفس فى بريطانيا
أفضل مقاييس العامل العام . ويتألف الاختبار من ٦٠ مصفوفة (أو رسماً)
حذف منها أحد أجزائها ، وعلى المحوص أن يحدد الجزء المحذوف بالانتقاء
من بين ٦ أو ٨ بدائل . وتتجمع الأسئلة فى ٥ سلاسل ، يحتوى كل منها على
١٢ مصفوفة متزايدة الصعوبة ولكنها متشابهة من حيث المبدأ . وتتطلب
المصفوفات الأولى دقة فى التمييز . أما المصفوفات المتأخرة والأكثر صعوبة

فتتضمن ادراك العلاقات المنطقية • وليس لهذا الاختبار زمن محدد (أى ينتمى الى فئة اختبارات القوة) • ويمكن تطبيقه فرديا أو جماعيا •

وقد نقلت هذا الاختبار الى اللغة العربية الهيئة الفنية للخدمة السيكولوجية العسكرية ، ولا تتوافر لدينا بيانات عن خصائصه السيكومترية للعسكريين، كما كتب الدارس العراقي قيس عبد الفتاح مهدي الى أحد مؤلفي هذا الكتاب (فؤاد أبو حطب) حول الدراسات التى أجريت حول هذا الاختبار بمركز البحوث التربوية والنفسية بجامعة بغداد وخاصة المصفوفات الملونة الذى تجرى محاولة تقنيه على المرحلة الابتدائية •

وقد تم تقنين الاختبار على البيئة السعودية عام ١٩٧٧ بواسطة فريق من الباحثين بإشراف فؤاد أبو حطب (٢٨) حيث طبق الاختبار على عينة تتألف من ٤٩٢٢ مفحوصا من البنين والبنات تمتد أعمارهم من ٨ سنوات وحتى ما بعد سن الثلاثين اختيروا من تلاميذ المرحلة الابتدائية والمتوسطة وطلاب المرحلة الثانوية والمرحلة الجامعية والخريجين بالمنطقة الغربية بالمملكة العربية السعودية • وحسب صدق التكوين الفرضي للاختبار ، حيث أكد صدق المحتوى أنه اختبار غير لغوي ، كما أكدت معاملات الارتباط بين درجات الاختبار والزمن المستغرق فى أدائه أنه اختبار قوة (حيث لم تكن هذه المعاملات دالة) • وثبتت أيضا صحة الترتيب التتابعى للمجموعات الخمس لأسئلة الاختبار باعتبارها متدرجة فى الصعوبة وذلك من خلال المقارنة بين متوسطات عينات الأعمار المختلفة فى كل مجموعة من المجموعات الخمس للأسئلة بالنسبة لكل درجة كلية يحصل عليها المفحوص • ثم تأكد أيضا أن تتابع المفردات داخل كل مجموعة من المجموعات الخمس منتظم من السهل الى الصعب وذلك من خلال فحص معاملات سهولة المفردات فى الأعمار المختلفة • كما استخدم فى حساب صدق التكوين الفرضى طريقة تمايز العمر وتمايز الصفوف الدراسية ، وقدّر كذلك الصدق المرتبط بالمعسكات باستخدام بعض الاختبارات العقلية الأخرى والتحصيل الدراسى وكانت معاملات الارتباط دالة •

وحسب الثبات بطريقة كيودر - ريتشاردسون وطريقة إعادة الاختبار وكانت المعاملات مرتفعة • وحسبت المعايير المثينة للاختبار •

(٢) اختبار رسم الرجل : أعدت هذا الاختبار الباحثة الأمريكية جود انف Goodenough عام ١٩٢٦ ، ويتطلب من المفحوص باختصار «رسم صورة لرجل» . وقد ظل هذا الاختبار في ميدان الاستخدام العام دون تعديل منذ ظهوره حتى عام ١٩٦٢ حيث ظهر تعديل جديد شامل له باسم «اختبار الرسم لجودانف وهاريس Goodenough-Harris Drawing Test . ويؤكد هذا التعديل كما يؤكد الاختبار الأصلي - دقة الطفل في الملاحظة وارتفاع تفكيره المجرد ، وليس المهارة الفنية في الرسم ، حيث تعطى درجة لكل جزء من الجسم رسمة الطفل ، وتفاصيل اللبس ، والنسب ، والمنظور ، وغير ذلك من الخصائص التي بلغت في طبعة ١٩٦٢ (٧٢ وحدة قابلة للقياس) . وقد اختيرت هذه الوحدات على أساس تمايز العمر ، والعلاقة بالدرجة الكلية في الاختبار ، والارتباط بدرجات اختبارات الذكاء الجماعية الأخرى . وتم الحصول على ذلك من اختبار ٥٠ طفلا ، و ٥٠ طفلة في كل مستوى دراسي ابتداء من الحضانة حتى سن ١٥ سنة . وقد أجريت دراسات كثيرة على المقياس المعدل ثبت منها أنه ثابت وصادق ، وأنه اختبار غير لفظي للنضج العقلي بالإضافة الى حساسيته الشديدة للعوامل الثقافية التي تؤثر في تمثيل الطفل لمفهوم شكل الانسان ، ومعنى ذلك أن هذا الاختبار ليس متحررا من أثر الثقافة كما كان مقترضا .

ويتطلب المقياس المعدل أن يرسم الطفل أيضا صورة لامرأة ، ويصح هذا المقياس الجديد الخاص برسم المرأة في ضوء ٧١ وحدة تتشابه مع مقياس رسم الرجل . كما يطلب من الطفل أيضا أن يرسم نفسه Self Scale ولو أن هذا المقياس الثالث أعد في جوهره كأسلوب استقاطي لدراسة الشخصية ، والبيانات المتوافرة عنه ليست مقنعة . وأعدت معايير لمقياس رسم الرجل ومقياس رسم المرأة من تطبيقه في أمريكا على ٣٠٠ طفل من كل مستوى عمري ابتداء من سن ٥ سنوات حتى سن ١٥ سنة . ويمكن تحويل الدرجات الخام في كل مقياس الى درجات معيارية معدلة متوسطها ١٠٠ وانحرافها ١٥ ، وحسب له الثبات بطريقة إعادة الاختبار وطريقته التجزئة النصفية . واستخدم في حساب صدقه محك تمايز العمر ومحك التنبؤ بالفجاء المدرسي ، كما استخدم في بعض الدراسات مقياس ستانفورد - بينيه كمحك أيضا ، وكذلك تقديرات المعلمين وتتابعات الدراسات السيكومترية حصول الاختبار في مختلف الثقافات نأكدت الدراسات الأمريكية اللاحقة جميعا ثبات (التقييم النفسي)

الاختبار وصدقه . وقد قنن الاختبار في البيئة المصرية وعدد من البيئات العربية الأخرى .

وقد تكون أولى الدراسات التي أجريت في هذا الصدد دراسة بإشراف القوصي (١٤٦) والتي قام فيها ميشيل اسكندر (١٤٦) بين عامي ١٩٢٤ ، ١٩٢٧ بتطبيق الاختبار على عينة مصرية اقتربت من الألفين ، أعمارهم بين ٤ ، ١٢ سنة ، ووضع نموذجا وسطا لكل سن كما أعطى معيارا مصرية للعمر العقلي ، وتحدد صدق القياس باستخدام محك اختبار الذكاء الابتدائي للقباني ، كما تمت المقارنة بين الصفوف الدراسية وبين الأعمار وبين الأطفال العاديين والمتخلفين وفي جميع الحالات كان التمييز في الاتجاه المتوقع على نحو يدل على صدق التكوين الفرضي . ولم يشر البحث الى حساب الثبات ، وبعد ذلك قام على علوي شلتوت بتعديل طريقة تصحيح الاختبار بما يلائم البيئة المصرية ، وظل الاختبار يستعمل في العيادة النفسية بكلية التربية جامعة عين شمس منذ ذلك الحين .

وقد قام مصطفى فهمي (٤٥) بدراسة على هذا الاختبار حيث طبق على ٩٧٦ تلميذا وتلميذة في فئات عمرية من ٦ سنوات حتى ١٢ سنة ، حسب من درجاتها معايير الأعمار العقلية مستخدما وحدة عمرية طولها سنة كاملة . ويلاحظ على هذا البحث بعض المشكلات المنهجية منها استخدام معادلة غير واضحة لحساب ثباته بطريقة كيودر - ريتشاردسون ، فقد أشار في مراجعته الى معادلة معينة واستخدم فيما يبدر معادلة أخرى ، بالإضافة الى الاعتماد على ٤٠ مفردة فقط ، بينما العدد الأصلي للمفردات - كما حددته جود انف في الاختبار الأصلي الذي اعتمد عليه في هذا البحث - هو ٥١ . كما لوحظ عدم وضوح محك تمايز العمر المستخدم في تقدير الصدق .

ومن الدراسات الهامة التي أجريت على هذا الاختبار تلك التي قام بها الباحث السوداني مالك البدرى (١٢٨ في ٢٧) وفيها لخص البحوث العربية السابقة التي أجريت حول الاختبار ومنها رسالة الماجستير التي تقدم بها محمد بطانية الى الجامعة الأمريكية في بيروت ، حيث طبق الاختبار على ٥٠٠ طفل أردني تراوحت أعمارهم بين ٥ و ١٠ سنوات . وحسب الثبات بطريقة إعادة الاختبار (ن = ٢٠) فبلغ ٨١٪ واستخدم طريقة تمايز العمر في

تقدير الصدق فوجد أن الفروق بين متوسطات الأطفال في مختلف الأعمار فروق دالة احصائية وفي اتجاه الزيادة مع التقدم في العمر . ولكن حين حسب معامل الارتباط بين درجات الاختبار وتقديرات المستوى الدراسي لم تكن لهذا المعامل دلالة احصائية واعتبر الباحث أن سبب هذه النتيجة إما عدم دقة المعلمين في تقدير مستوى التلاميذ أو عدم اعتماد التحصيل على الذكاء وحده . وغاب عن الباحث تفسير ثالث قد يكون أكثر صحة وهو عدم ارتباط الاختبار بمعك التحصيل . وهي نتيجة تتفق مع معظم النتائج التي توصل إليها الباحثون في عدد من الثقافات الأخرى . واستخرج محمد بطانيه معايير مؤقتة لذكاء الطفل الأردني .

وفي العراق قام عبد الجليل الزوبعي (٢٧ : ٦٥) بتطبيق الاختبار على عدد من الأطفال وقارن بين نسب الذكاء حسب تعليمات جودائف والتحصيل المدرسي فلم يجد مرة أخرى ارتباطا عاليا ، واستنتج الباحث من ذلك الاستنتاج الذي أشرنا إليه في الفقرة السابقة ، وهو احتمال أن هذا الاختبار لا يصلح لقياس ذكاء الطفل العراقي .

وفي السودان قام مالك البدرى (٢٧) بتطبيق الاختبار على ١٤٢٥ طفلا وطفلة من سن ما قبل المدرسة والمرحلة الابتدائية في الريف والحضر السودانيين ، وصحح الاختبار تبعاً لفتح جودائف الأصلي (٥١ مفردة) . ولتقدير الصدق حسب معامل الارتباط بين الاختبار والمستوى الدراسي فلم يجد الارتباط دالا احصائيا مرة ثالثة . وفي ذلك يقول الباحث بأن الاختبار بصورته الحالية ، قد لا يكون صالحا لقياس ذكاء الطفل العربي والافريقي . وأن العمل الأساسي للاخصائيين النفسيين العرب في هذا المجال ليس هو تطبيق الاختبار واستخراج المعايير فحسب ، بل إعادة النظر في اختبار رسم الرجل بشكله الحالي ومحاولة تعديله حتى يلائم الطفل العربي والافريقي في بيئاته المختلفة ، (٢٧ : ٧٠) . وقد استطلع الباحث مجموعة من الفروض حول صدق الاختبار قد يكون أهمها أن من أسباب عاقبة تطبيق اختبار رسم الرجل في البيئات العربية والافريقية الاختلافات الكبيرة لأزياء الرجل لو قررن بزي الرجل الأوربي والأمريكي الأكثر ثباتا . وتأكدت صحة هذا الفرض من مقارنة متوسط درجات من يرسمون رجلا بالملابس الأوربية بمتوسط من يرسمون رجلا بالملابس التقليدية . ولذلك قام الباحث بفحص

مجموعة من رسوم الأطفال الذين يرسمون الرجل بالملابس التقليدية ، واقتراح بعض التعديلات الجوهرية على مفردات الاختيار الأصلي وإضافة بعض العناصر الجديدة لإعطاء من يرسمون الرجل بالملابس التقليدية فرصة متساوية مع أولئك الذين يرسمون رجلا بالملابس الأوربية . ومن أهم العناصر التي اقترح الباحث إضافتها : رسم الشارب أو اللحية أو كليهما ، إظهار الوشم ، تظليل الجسم ، رسم العصا في اليد . ويبدو أن نتائج مالك البدرى تصلح لكثير من الدول العربية التي يرتدى الرجال فيها « الجلابية » أو « الدشداشة » أو « الثوب » كالسعودية والكويت ومنطقة الخليج وبعض أنحاء مصر والمغرب العربي .

وفي الكويت أجرى محمد نسيم رافت (١٤٤) دراسة على الاختبار تهدف إلى تحديد نوعين من المعايير أحدهما يناسب الزى الأوربي وثانيهما الزى التقليدي الكويتي وقد أجريت الدراسة على ٢٥٦٢ طفلا وطفلة تمتد أعمارهم بين ٤ ، ١٢ سنة .

وفي لبنان قام عبد السلام عبد الغفار والسيد فؤاد أعظمى بتطبيق الاختبار على عينة مكونة من ١٨٥٥ طفلا وطفلة تتراوح أعمارهم بين ٦ ، ١٢ سنة واستخدم في تصحيح الاختبار ٥٠ مفردة فقط من مفتاح جودائف الأصلي ، وحسب الثبات بطريقة إعادة الاختبار فبلغ الارتباط ٧٥ر ، وفي حساب الصدق استخدم محك تمايز العمر ، وأعدت معايير أعمار عقلية .

وفي اليمن قام فتحى السيد عبد الرحيم (١٢٩) بتطبيق الاختبار على أطفال منطقة صنعاء من ذوى الأعمار الزمنية التي تمتد من ٦ - ١٢ سنة بلغ عددهم ١٤٨٢ طفلا وطفلة . وحذا الباحث في بحثه حذو بحث مصطفى فهمى السابق الإشارة إليه ، فأستخدم مفرداته الأربعين وحسب معايير العمر لوحدة زمنية طولها عام كامل واستخدم طريقة كيبودر - ريتشاردسون في حساب الثبات . وفي حساب الصدق استخدم محك تمايز فكانت الفروق دالة بين الأعمار ، كما حسب معامل الارتباط بين الاختبار ومحك تقديرات المعلمين فبلغ ٥٩ر وهو معامل مرتفع نسبيا بالمقارنة بالمعاملات التي لوحظت في البحوث السابق الإشارة إليها .

وقد أجريت هذه الدراسات على اختبار جودائف الأصل طبعة ١٩٢٦ ،
أما بالنسبة للاختبار المعدل لجودائف وهاريس الصادر عام ١٩٦٢ فلم يجر
عليه إلا عدد محدود من الدراسات في الثقافتين المصرية والعربية . فقد
قام محمد فرغلي قراج وعبد الحليم محمود السيد وصفية مجدى
بترجمة كراسة تعليمات هذا الاختبار عام ١٩٧٦ ، وفي نفس العام أجرى
محمد متولى غنيمه (٤٤) تقنين هذا المقياس الجديد على البيئة المصرية
حيث طبق الاختبار على ١٩١٥ طفلا وطفلة من سبع مدارس ابتدائية بمحافظة
القاهرة . وعلى الرغم من أن بعض الأطفال رسموا الرجل مرتديا الجلباب إلا
أن الباحث اعتمد في التحليل على أولئك الذين رسموا الرجل مرتديا الزي
الأوربي مستخدما مفتاح هاريس المعدل الذي يتألف من ٧٢ مفردة ، وحسب
الثبات بطريقة تكرار المصحح الواحد ، وتكرار المصححين المختلفين ، فبلغ
معامل الموضوعية في الحالة الأولى ٩٣ ر ، وفي الحالة الثانية بين ٦١ ر ،
٨٧ ر ، وكذلك استخدم طريقة إعادة الاختبار فبلغ معامل الثبات ٨٤ ر ، ٩٨ ر
للاعمار المختلفة . وفي حساب الصدق المرتبط بالمحكات استخدم الباحث
مقياس ستانفرد - بينيه وبلغ معامل الارتباط بين ٨٠ ر ، ٥٤ ر ، والنحصيل
الدراسي وكانت معاملات الارتباط بين ٤٢ ر ، ٨١ ر . كما قارن الباحث بين
مجموعتين متضادتين أحدهما من المتخلفين عقليا والثانية من الأطفال
العاديين وكانت الفروق دالة في الاتجاه المتوقع ، كما استخدم الباحث محك
تمايز العمر ، ولم يظهر البحث فروقا دالة بين الجنسين في معظم الصفوف
الدراسية .

ولعل من المشكلات الجوهرية التي يظهرها هذا الاختبار مسألة زى
الرجل المرسوم ، وهي مشكلة حاول أن يحلها كل من مالك البدرى ومحمد
متولى غنيمه عن طريق التضمين أى التساهل في تصحيح بعض المفردات .
وقد قامت أمال أحمد مختار صادق (١٩) في عام ١٩٧٩ بدراسة مقارنة
مستقلة للمقارنة بين رسوم الرجل مرتديا الزي التقليدى السعودى والزي
الأوربي ، وقد أجريت الدراسة على عينة من الأطفال قوامها ١٩٢٢ طفلا
وطفلة تمتد أعمارهم بين ٨ سنوات ، ١٥ سنة . وقد اختارت سن ٨ سنوات
كحد أدنى لأنه العمر الذي أظهرت البيانات الامبريقية أن نوعى الزي يتمايزان
فيه في رسوم الأطفال ، وقامت الباحثة بدراسة الفروق بين النسب لكل مفردة

من مفردات المقياس في نوعي الرسم وكانت الفروق الدالة جميعا لصالح
الزى الأوربي ، وتعلق الباحث على هذه النتيجة بقولها (١٩ : ١١٩) :

« تؤكد هذه النتيجة أن رسم الرجل بالزى الأوربي يهيء الفرصة للطفل
أن يتنبه إلى تفاصيل أكثر لا ترجع إلى طبيعة الزى وحده .. فكثير من
المفردات التي أظهرت فروقا دالة لصالح الزى الأوربي تتضمن تفاصيل الجسم
البشرى ذاته. ومن ذلك وجود الرقبة ووجود الشعر » .

ومن الدراسات التقنية الهامة التي أجريت على اختبار جودائف -
هاريس طبعة ١٩٦٣ تلك التي أجريت في المملكة العربية السعودية عام
١٩٧٧ وقام بها فريق من الباحثين بإشراف فؤاد أبو حطب (٢٧) ، ولعل
أهم الإضافات في هذا البحث أن بروتوكولات التصحيح المستخدمة (وهي
من اعداد فؤاد أبو حطب) جعلت عدد المفردات ٧٧ مفردة نتيجة إضافة
مفردات جديدة أو مفردات تضمنها الاختبار الأصلي (طبعة ١٩٢٦) . وقد
أجرى البحث على ٢١٥ طفلا وطفلة من مرحلة ما قبل المدرسة تمتد أعمارهم
بين ٢ - ٥ سنوات ، ١٩١١ طفلا وطفلة تمتد أعمارهم بين ٦ - ١٥ سنة
اختبروا جميعا من رياض الأطفال والمدارس الابتدائية والمتوسطة بالمنطقة
الغربية بالمملكة العربية السعودية . وحسبت معاملات شيوخ المفردات التي
أظهرت أن عدد المفردات التي تعد أفضل مقياس للذكاء باستخدام رسم الرجل
مرتديا ثوبا (الزى الوطني السعودي) هو ٦٦ مفردة ، أما باقي المفردات
فهى أكثر ملاءمة لرسم الرجل مرتديا الزى الأوربي .

وفي حساب صدق المفردات استخدمت طريقة المقارنة الطرفية وحسبت
معاملات الارتباط الثنائية بين المفردة والدرجة الكلية في الاختبار . أما
بالنسبة لصدق الاختبار ككل استخدم محك تمايز العمر في تقدير صدق
التكوين الفرضي ، والأداء في اختبار المصفوفات المتتابعة (والذى سبق تقنيته
على نفس الثقافة) لتقدير الصدق المرتبط بالمحكات . أما عن ثبات الاختبار
فقد حسب بطريقة إعادة الاختبار وبطريقة كيودر - ريتشاردسون ، وكانت
المعاملات - كما هو متوقع - منخفضة في الأعمار الدنيا - وخاصة أعمار ما
قبل المدرسة ، أما في الأعمار الأعلى فكانت مرتفعة نسبيا ، وكانت بالنسبة
لطريقة إعادة الاختبار أقل منها بالنسبة لطريقة كيودر - ريتشاردسون ، وهي

نتيجة متوقعة في ضوء طبيعة كل من الطريقتين في حساب الثبات ، وتباين الخطأ المتضمن في كل منهما .



وتحتل مسألة الموضوعية أهمية خاصة في اختبار رسم الرجل ، وقد حسبت بالطرق الثلاث الآتية :

١ - إعادة المصحح الواحد باستخدام بروتوكول التصحيح في إحدى المرتين وبدون استخدامها في المرة الثانية .

٢ - إعادة المصححين باستخدام بروتوكولات التصحيح بالنسبة لبعض المصححين وبدون استخدامها بالنسبة للبعض الآخر .

٣ - إعادة المصحح الواحد باستخدام بروتوكولات التصحيح في الحالتين .

وكانت معاملات الموضوعية مرتفعة ، إلا أنها كانت منخفضة في حالة الاختلاف في استخدام البروتوكولات أو عدم استخدامها .

وحسبت للاختبار معايير في صورة نسب ذكاء انحرافية متوسطة ١٠٠ وانحرافها المعياري ١٥ لجميع الأعمار .

في عام ١٩٨٢ قام نعيم عطية (٤٧) بمحاولة بناء نسق نظري للاختبار له طبيعة هرمية يتألف من عدة عمليات عقلية يحتاجها الأداء في الاختبار وتشمل : ثبات الصورة الحسية ، تركيز مواقعها ، ادراك الحجم النسبي ، ادراك التناسب الداخلي ، ادراك أبعاد التجسيم ، الاتساق البصري - الحركي ، توجيه الخطوط ، الإيماء والتعبير ، الاستكمال ، الرسم الجانبي . ويقترح ثلاث مراحل لإجراء الاختبار وهي :

١ - الرسم المواجه ، وفيها يطلب من الطفل رسم صورتين للرجل في وضع مواجه أحدهما تلقائية والثانية في ضوء تنبيهات معينة .

٢ - استقصاء الرسوم واستكمالها وتآلف من خطرتين : الاستقصاء لاستدراج التسميات والاستقصاء لاستكمال النواقص .

٣ - الرسم الجانبي .

ويصف الباحث بالتفصيل بروتوكولات التصحيح المقترحة لعدد من المفردات بلغ ٢٢٦ مفردة . ولم يجز الباحث تقنيا شاملا للمقياس بسبب ظروف لبنان - وطنه ، واكتفى بدراسة امكانات التعديل الذي يقترحه للتقنين بأن طبقه على عينة صغيرة نسبيا (ن = ٢٢٩) من أطفال تمتد اعمارهم بين ٤ و ٨ سنة ، ٨ سنوات . وحسب الثبات بأعادة الاختبار وبلغ معامل الارتباط ٧١ر (ن = ٥٥) ، واستخدم في حساب الصدق محك تمايز العمر (أو ما يسميه الباحث تدرج النتائج مع السن) . أما المعايير فأتخذت صورة نسب الذكاء التقليدية بمتوسط ١٠٠ وانحراف معياري ١٦ .

٣ - اختبار هايدلبرج للذكاء :

هذا الاختبار أعد في أصله الألماني كراتزيمر Kratzmeier عام ١٩٧٧ ونقله الى العربية حسنين الكامل (١٦) عام ١٩٨٥ ، ويقوم في جوهره على تحليل ما يسمى ، القدرة التمييزية ، الى خمسة مكونات هي : ادراك العلاقات ، والسلاسل المنطقية ، والانتاجية والتمييز والابتكارية . وقد استبعد معد الاختبار بالعربية القسم الخامس (الابتكارية) من الصورة العربية . وطبقه على عينة مكونة من ٢١٦ طفلا وطفلة من تلاميذ الصفوف الثاني والثالث والرابع من مرحلة التعليم الأساسي بمحافظة سوهاج ، وأجرى على كل اختبار فرعي (يتألف من ٧٢ مفردة) تحليلا للمفردات بحساب معاملات الصعوبة والتمييز ثم طبق الاختبار على عينة أخرى (ن = ٨٢٤) لها نفس الخصائص السابقة وحسبت لها المعايير في صورة مئينيات ثم حسبت نسب الذكاء الانحرافية ، بمتوسط ١٠٠ وانحراف معياري ١٥ والدرجات الثائية .

وحسب الثبات بمعامل الفا لكرونيك . أما صدق التكوين الفرضي فقد حسب بطريقتين : اولاهما تمايز العمر ، وكانت الفروق في الاتجاه المتوقع ،

والطريقة الثانية هي التحليل العاملي وتوصل الباحث الى اربعة عوامل تطابقت مع الاختبارات الفرعية فيه . وحسب الصديق القلازمي باستخدام محك التحصيل المدرسي وكانت معظم معاملات الارتباط دالة ، وباستخدام اختبار الذكاء المصور كمحك ، وكانت معاملات الارتباط منخفضة بشكل واضح يوحي بضرورة اعادة تقنين اختبار الذكاء المصور ذاته واعادة النظر فيه كمقياس للذكاء .

مراجع الباب الثالث

- ١ - أحمد زكى صالح : الأسس النفسية للتعليم الثانوى • دار النهضة العربية ، ١٩٧٢ •
- ٢ - أحمد زكى صالح : علم النفس التربوى • مكتبة النهضة المصرية ، ١٩٧٢ •
- ٣ - أحمد زكى صالح : علم النفس فى الإدارة والصناعة • مكتبة النهضة المصرية ، ١٩٧٢ •
- ٤ - أحمد زكى صالح : تعليمات اختبار الذكاء المصور • (ب • ت)
- ٥ - أحمد فكرى : الطريقة المبتكرة لقياس العقول ، ١٩٢٠ •
- ٦ - اسماعيل القبانى : مقياس ستانفورد - بينيه للذكاء ، لجنة لتأليف والترجمة والنشر ، ١٩٢٨ •
- ٧ - اسماعيل القبانى : قياس الذكاء • المطبعة الأميرية ، ١٩٢٨
- ٨ - اسماعيل القبانى : اختبار الذكاء الابتدائى ، لجنة التأليف والترجمة والنشر (ط ثانية) ، ١٩٣٩ •
- ٩ - امان أحمد مختار صادق • الصورة السعودية لمقياس ستانفورد - بينيه (بحث غير منشور) ، ١٩٧٧ •
- ١٩ - _____ : دراسة مقارنة لاختبار رسم الرجل فى كل من الزى التقليدى والأوروبى • فى بحوث فى تقنين الاختبارات النفسية (تحرير فؤاد أبو حطب) ، المجلد الثانى ، مكتبة الأنجلو المصرية ، ١٩٧٩ ، ١٠٥ - ١٢٨ •
- ١٠ - أمير بقطر : اختبارات الذكاء ، مجلة التربية الحديثة ، ١٩٢٨ •
- ١١ - جابر عبد الحميد جابر ، يوسف الشيخ : علم النفس الصناعى • دار النهضة العربية ، ١٩٦٨ •

١٢ - جابر عبد الحميد جابر: الذكاء ومقاييسه • دار النهضة العربية
١٩٧١ •

١٣ - جمال محمد علي : تقنين مقياس ستانفورد - بينيه للذكاء على تلاميذ
مرحلة التعليم الأساسي • رسالة ماجستير ، كلية التربية جامعة
عين شمس ، ١٩٨٢ •

١٤ - حسن حسين زيتون : اختبار مراحل بياجيه للنمو العقلي • الكويت:
دار القلم ، ١٩٨٦ •

١٥ - حسن عمر : مقياس الذكاء ، ١٩٢٨

١٦ - حسنين الكامل : اختبار هايدلبرج للذكاء • مكتبة النهضة المصرية ،
١٩٨٥

١١٦ - رجاء محمود علام ، كمال ابراهيم مرسى : مقياس ستانفورد - بينيه
للذكاء ، الصورة الكويتية ، ادارة الخدمة النفسية لوزارة التربية ،
الكويت ، ١٩٧٨ •

١٧ - سميرة سعيد عبد القادر : دراسة تجريبية لمعامل التدهور في ذكاء
الذهانيين كما يبينه مقياس وكسلر - بلفير • رسالة ماجستير ، كلية
الآداب جامعة عين شمس ، ١٩٧٧ •

١٨ - سيد أحمد عثمان ، فؤاد أبو حطب : التفكير ، دراسات نفسية •
الأنجلو المصرية ، (ط ١) ١٩٧٢ ، ط ٢ ، ١٩٧٧ •

٢٠ - السيد محمد خيرى : تعليمات اختبار الذكاء الاعدادى • دار النهضة
العربية (ب • ت) •

٢١ - السيد محمد خيرى : تعليمات اختبار الذكاء العالى • دار النهضة
العربية (ب • ت) •

٢٢ - صفوت فرج : القياس النفسى • دار الفكر العربى ، ١٩٨٠

١٢٢ - عبد العزيز القرصى ، هدى بريدة ، حامد زهران : اختبار عين شمس
للذكاء الابتدائى • عالم الكتب ، ١٩٧٤

- ٢٣ - عبد الله عبد العزيز المعيلي : أساليب تقنين الاختبار وتطبيق ذلك على اختبار ستانفورد - بينيه ليناسب طلبة المدارس الابتدائية بالرياض رسالة ماجستير ، كلية التربية جامعة الملك سعود ، ١٤٠٢ هـ .
- ٢٤ - فؤاد أبو حطب : القدرات العقلية . الأنجلو المصرية ، (ط ٤) ١٩٨٣
- ٢٥ - فؤاد أبو حطب (تحرير) بحوث في تقنين الاختبارات النفسية (المجلد الأول) . مكتبة الأنجلو المصرية ، ١٩٧٧ .
- ٢٦ - _____ : (تحرير) بحوث في تقنين الاختبارات النفسية (المجلد الثاني) . مكتبة الأنجلو المصرية ، ١٩٧٩ .
- ٢٧ - _____ وآخرون : تقنين اختبار رسم الرجل على البيئة السعودية ، في المرجع ٢٥ ، ٢٤٧ - ٢٤٦
- ٢٨ - _____ وآخرون : تقنين اختبار المصنفات المتتابعة في المرجع ٢٥ ، ١٩١ - ٢٤٦
- ٢٩ - فؤاد البهي السيد : الذكاء . دار الفكر العربي ، ١٩٧١
- ١٢٩ - فتحي السيد عبد الرحيم : اختبار رسم الرجل ، تطبيقه وتقنيته على البيئة اليمنية . بحث غير منشور ، ١٩٧٥
- ٣٠ - فرج عبد القادر طه : العلاقة بين الاصابات في الصناعة وكل من الصفحة النفسية للذكاء والسرعة الاراكية والسرعة الحركية . المجلة الاجتماعية القومية - المجلد السادس ، سبتمبر ١٩٦٩ ، ٧٣ - ١٠٤
- ٣١ - _____ ، صلاح أحمد مرحاب : الصورة المغربية لمقياس ركنسلر - بلفيو لذكاء الراشدين والمراهقين المغرب : الرباط ، مطبعة الكوثر ، ١٩٧٧

٢٢ - فرس ، برايان (ترجمة فؤاد أبو حطب) - افاق جديدة في علم النفس
عالم الكتب ، ١٩٧٢

٢٣ - لويس كامل مليكة : الدلالات الاكلينيكية لمقياس وكسلر - بلفيو لذكاء
الراشدين والمراهقين - النهضة المصرية ، ١٩٦٠

٢٤ - لويس كامل مليكة : المقدرات في قياس الذكاء - النهضة المصرية ،
١٩٦٠

٢٥ - لويس كامل مليكة : نماذج التصحيح وجداول الدرجات الموزونة
ونسب الذكاء لمقياس وكسلر بلفيو لذكاء الراشدين والمراهقين :
النهضة المصرية ، ١٩٦٠

٢٦ - مائسة انور المفتي : اعداد صورة معدلة لاختبار وكسلر - بلفيو
لذكاء الراشدين والمراهقين - رسالة ، ماجستير ، كلية الاداب جامعة
عين شمس ، ١٩٧١

٢٧ - مالك البدرى : سيكولوجية رسوم الاطفال : اختبارات رسم الانسان
وتطبيقاتها على اطفال البلاد العربية - بيروت : دار الفتح للطباعة
والنشر ، ١٩٦٦

٢٨ ١ - محمد بطانية : تقنين اختبار رسم الرجل على البيئة الاردنية -
رسالة ماجستير ، الجامعة الاردنية ، في المرجع رقم ٢٧

٢٨ - ماهر محمود الهوارى : دراسة لصدق مقياس ستانفورد - بينيه للذكاء
رسالة ماجستير غير منشورة - كلية الاداب جامعة عين شمس ، ١٩٦٣

٢٩ - محمد خليفة بركات : الاختبارات والمقاييس العقلية - مكتبة مصر ،
١٩٥٤

٤٠ - محمد عبد السلام أحمد ، لويس كامل مليكة : مقياس ستانفورد -
بينيه للذكاء - مكتبة النهضة المصرية ، ١٩٥٦

٤١ - محمد عماد الدين اسماعيل ، لويس كامل مليكة : مقياس وكسلر -
بلفيو لذكاء الراشدين والمراهقين - مكتبة النهضة المصرية ، ١٩٥٦

٤٢ - محمد عمار الدين اسماعيل • لويس كامل مليكة : مقياس وكسلر
لذكاء الأطفال • مكتبة النهضة المصرية ، ١٩٦١ •

٤٣ - محمد عمار الدين اسماعيل : بحث في اختبار وكسلر لذكاء الراشدين
والمراهقين • مكتبة النهضة المصرية (ب • ت)

٤٤ - محمد متولى أحمد غنيم : تقنين اختبار رسم الرجل بالنسبة لتلاميذ
المرحلة الابتدائية في مصر • رسالة ماجستير ، كلية التربية جامعة
عين شمس ، ١٩٧٦ •

١٤٤ - محمد نسيم رافت : تقنين اختبار رسم الرجل على البيئة الكويتية .
الكويت ، التقرير السنوى لكتب الخدمة الاجتماعية ، ادارة التربية ،
وزارة التربية بالكويت ، ١٩٦٨

٤٥ - مصطفى فهمى : اختبار رسم الرجل • دار مصر للطباعة ، (ب • ت)

٤٦ - المقتطف : مقياس الذكاء ، أغسطس ١٩١٧

١٤٦ - ميشيل أسكندر : رسوم الأطفال وعلاقتها بنموهم العقلى • رسالة
ماجستير ، المعهد العالى للتربية للمعلمين (كلية التربية جامعة عين
شمس) ، ١٩٣٥

٤٧ - نعيم عطية : ذكاء الأطفال من خلال الرسوم ، نسق جديد لاختبار
رسم الرجل ، دراسة استطلاعية ، بيروت : دار الطليعة ، ١٩٨٢

٤٨ - هاريس ، د • ب (ترجمة محمد فرغلى فراج ، عبد الحليم محمود
السيد ، صفية مجدى) ، اختبار الرسم لجودانف وهاريس (غير
منشور) ، ١٩٧٦

49. Anastasi, A. Psychological Testing, Macmillan, 1961, 1958, 1976, 1982.
50. Binet, A., Simon, T. The development of intelligence in children, In : Jenkins, J.J. and Patterson, D.G. (ed.) Studies in Individual Differences. London : Methuen, 1961.

51. Bond, E.A. Tenth Grade Abilities and Achievements. Teacher College, 1940.
52. Bradway, K.P., Thompson, C.W., Grays, R.B. Preschool IQ's after twenty-five years. J. Educ., Psychol., 1958, 49, 278-181.
53. Burt, C. Mental and Scholastic Tests. (4th ed.) London, Staples Press, 1962.
54. Carlton, J.S. and Jensen, C.M. The factorial structure of the Raven Coloured progressive Matrices. Educ. and Psychol. Mesment, 1980, 40, 111-1115.
55. Cohen, J.A. factor-analytically based rational for the Wechsler Adult Intelligence Scale. J. Consul. Psychol. 1957, 21, 451, 457.
56. Cohen, J. The factorial structure of the WAIS between early adulthood and old age. J. Consul. Psychol., 1957, 283-390.
57. Cronbach, L.J. Essentials of Psychological Testing. New York, Harper, Brothers, 1960, 1970, 1984.
58. Edwards, A.J. Individual mental testing, vol. I, Scraton : Intext Educ. Publ., 1971.
59. ——— Individual mental testing, vol. II, Scraton. Intext Educ. Publ., 1972.
60. Freeman, F.S. Theory and practice of psychological testing. (3rd. ed.), New Delhi Oxford, 1972.
61. Guertin, W.H., et al. Research with the Wechster Scales for Adults: 1960-1965 Psychol. Bull., 1966, 66, 385-409.
62. ——— et al. Research with Wechsler Scales for Adults: 1965-1970 Psychol. Rec., 1971, 21, 289-339.

63. Guilford, J.P. *The Nature of Human Intelligence*. McGraw-Hill, 1967.
64. Himmelstein, P. Research with the Stanford-Binet, Form L.M. The first five years. *Psychol. Bull.*, 65, 156-164, 1966.
65. Jones, L.V. A factor analysis of the Stanford-Binet at four age levels. *Psychometrika*, 1949, 14, 299-331.
66. Jones, L.V. Primary abilities in the Stanford Binet, age 13. *J. Gen. Psychol.*, 1954, 74, 126-147.
67. Kennedy, W.A., et al. The ceiling of the new Stanford-Binet. *J. Clin. Psychol.*, 1960, 17, 274-276.
68. Matarazzo, J.D. *Wechsler's measurement and appraisal of adult intelligence* (5th ed.) Baltimore: Williams and Wilkins, 1972.
69. McNemar, Q. *The revision of the Stanford-Binet Scale : An analysis of the Standardization data*. Boston : Houghton Mifflin, 1942.
70. Moriarty A.E. children's way of coping with the intelligence test. *Menninger Clinic Bull.*, 1960, 24, 115-127.
71. Peterson, J. *Early Conceptions and tests of intelligence*. New York, Yonkers, Wold Book, 1926.
72. Pineau, S.R. *Changes in intelligence quotient from infancy to maturity*. Boston : Houghton Mifflin, 1961.
73. Sattler, J.M. *Assessment of children's intelligence and special abilities* (2nd ed.), Boston : Allyn and Bacon, 1972.
74. Saunders, D.R. On the dimensionality of the WAIS battery for two groups of normal males. *Psychol. Rep.*, 1959, 5, 529-541.

75. Saunders, D.R. A factor analysis of the information and Arithmetic items of the WAIS, *Psychol. Rep.*, 1960, 6, 367-383.
76. Saunders, D.R. A factor analysis of the Picture Completion items of the WAIS. *J. Clin. Psychol.*, 1960, 16, 146-149.
77. Saunders, D.R. Alpha frequency : A cross-validation. *J. Cli. Psychol.*, 1961, 17, 165-167.
78. Terman, L.M., Merrill, M.A. *Measurement of intelligence*, Boston, Houghton Mifflin, 1937.
79. Terman, L.M., Merrill, M.A. *Stanford-Binet Intelligence Scale. Manual for the third revision, Form L-M*. Boston : Houghton Mifflin, 1960.
80. Terman, L.M. and Merrill, M.A. *Stanford-Binet Intelligence Scale : 1972 norms edition*, Boston : Houghton Mifflin, 1973.
81. Thorndike, R.L., Hagen, E.P. and Sattler, J.M. *Stanford-Binet Intelligence Scale, (4th ed.)* Chicago : The Riverside Publishing Co., 1985.
82. Vernon, P.E. *Intelligence and attainment tests*. University of London Press, 1960.
83. Wechsler, . *The measurement and appraisal of adult intelligence (4th ed.)* Baltimore : Williams and Wilkins, 1958.
84. Wolf, T.H. *Alfred Binet*. Chicago : Univ. of Chicago Press, 1973.

الباب الرابع

مقاييس الاستعدادات

الفصل التاسع

بطاريات الاستعدادات المتعددة

يقصد ببطاريات الاستعدادات المتعددة multiple-aptitude batteries مجموعة المقاييس التي تهدف الى قياس مجموعة من مظاهر النشاط العقلي، كل على حدة . قياسا مستقلا ، وتعطى لكل منها درجة مستقلة . وعادة ما لا تعطى هذه البطاريات مؤشرا واحدا للبطارية ككل (مثل نسبة الذكاء أو العمر العقلي أو الدرجة المعيارية) . وفي هذا تختلف بطاريات الاستعدادات المتعددة عن اختبارات الذكاء العام التي تناولناها في الباب السابق ، كما تختلف عنها أيضا في أن اختبارات الذكاء العام تعتمد في بنائها على درجة تجانس أسئلتها أو اختبارات الفرعية بحيث تعطى في النهاية مقياسا للذكاء يتوافق فيه شرط الاتساق الداخلي ، حتى ولو كانت هذه الاختبارات تعطى درجات منفصلة للاقسام اللفظية أو العملية فيها (مثل مقاييس وكسلر) أو للاقسام اللفظية والعديدية (مثل اختبارات ذكاء تلاميذ المرحلة الثانوية وطلاب الجامعات) ، بالإضافة الى الدرجة الكلية أو التقدير العام للذكاء . أما في بطاريات الاستعدادات المتعددة فإن الهدف من القياس هو إعطاء الفحوص عددا من الدرجات المنفصلة للتمييز بين القدرات العديدة التي تظهر اختلافات داخل الفرد الواحد بقدر ما تظهرها بين الأفراد .

ويمكن أن نلخص أهم العوامل التي أدت الى زيادة الاهتمام بهذا الاتجاه « الفارق » وخاصة منذ نهاية الحرب العالمية الثانية فيما يلي :

- ١ - زيادة الاعتراف بالفروق الفردية داخل الفرد intraindividual variations في أدائه على الأجزاء المختلفة لاختبارات الذكاء العام - سواء كانت أسئلة منفصلة أو اختبارات فرعية . وقد تناول أحد مؤلفي هذا الكتاب في موضع آخر (١٩) الأدلة والشواهد التجريبية على وجود هذه الفروق داخل الفرد الواحد .

٢ - زيادة الاعتقاد بأن ما يسمى اختبارات الذكاء العام أقل عمومية مما كان مفترضا فيها . وتوجد شواهد كثيرة على أن بعض هذه الاختبارات ، كمقياس ستانفورد - بينيه ، يقيس الفهم اللفظي بالإضافة الى عدد من القدرات الأخرى .

٣ - لم تتناول اختبارات الذكاء العام بعض جوانب النشاط العقلي مثل القدرات الميكانيكية والموسيقية والكتابية ، باستثناء قليل من الاختبارات العملية للذكاء . ولذلك سرعان ما ظهرت اختبارات « الاستعدادات الخاصة ، special aptitudes » - كما تسمى ولا زالت في بعض المؤلفات المتخصصة - لتكمل نقائص اختبارات الذكاء العام .

٤ - زيادة نشاط علماء النفس في المجالات التطبيقية - كالتوجيه والإرشاد التربوي والمهني ، والانتقاء والتصنيف في الميادين الصناعية والعسكرية - مما تطلب توافر الاختبارات التي تقيس الاستعدادات المتعددة ، بل أن ظهور بعض الاختبارات المبكرة التي تقيس « الاستعدادات الخاصة » ، مثل الاستعداد الكتابي والميكانيكي ، كان انعكاسا لهذا الاهتمام .

٥ - التوسع في استخدام منهج التحليل العاملي عند بناء الاختبارات العقلية . وقد أمكن لعلماء النفس بهذا المنهج أن يتوصلوا الى تحديد أدق للقدرات المختلفة التي يتضمنها ذلك المصطلح العام « الذكاء » ، بالإضافة الى تصنيف هذه القدرات والوصول الى نماذج نظرية لما يسمى بالتنظيم العقلي (راجع ١٩) ولذلك ظهرت البطاريات التي تقيس « عوامل متعددة » ، ويتطلب هذا بالطبع أن تتألف من مجموعة من الاختبارات ، يقيس كل منها عاملا من العوامل التي أمكن تحديدها بالتحليل العاملي .

وسوف نتناول في هذا الفصل طائفة من هذه البطاريات التي تقيس عددا من الاستعدادات في وقت واحد وتسمى لهذا السبب بطاريات الاستعدادات المتعددة .

اختبارات القدرات العقلية الأولية :

تعد بطارية اختبارات القدرات العقلية الأولية من أهم نتائج استخدام

منهج التحليل العاملي في بناء الاختبارات . وقد أعد هذه البطارية العالم الأمريكي ثurstone محصلة لنشاطه العلمي الواسع في ميدان تحليل القدرات العقلية . وقد نشرت في عام ١٩٤١ باسم : اختبارات شيكاغو للقدرات العقلية الأولية The Chicago Tests of Primary Mental Abilities وظلت تصدر بهذا الاسم حتى عام ١٩٤٧ . وفيها انتقى ثurstون لكل عامل الاختبارات التي تشبعت به تشبعا عاليا ، وأعدت لطلاب المرحلة الثانوية والجامعات . وبلغ عدد الاختبارات ١١ اختبارا تقيس ٦ عوامل أولية هي : الفهم اللفظي (V) ، والقدرة المكانية (S) والقدرة العددية (N) والذاكرة (M) ، وطلاقة الكلمات (W) ، والاستدلال (R) . واستخدام لقياس كل قدرة اختباران ، ما عدا الذاكرة فقد استخدم في قياسها اختبار واحد ، وقد طبعت هذه الاختبارات في كراسات بحيث يمكن تطبيقها في حصتين دراسيتين (أى في حوالى ساعتين) .

ثم أدخلت تعديلات على هذه البطارية منذ عام ١٩٤٧ فأصبحت تسمى اختبارات SRA للقدرات العقلية الأولية نسبة الى الهيئة العلمية التي تصدرها .

ومن أهم هذه التعديلات اختزال طولها وخفض مستواها الى المستويات العمرية الأدنى . وأصبح لها ٣ مستويات . من ٥ - ٧ سنوات ، ومن ٧ - ١١ سنة ، ومن ١١ - ١٧ سنة . وقد تعرضت هذه البطارية (طبعة ١٩٤٧) للنقد في معظم مجلدات الكتاب السنوى للقياس العقلي الذي يصدره بورس ، بالإضافة الى انتقادات سوبر لها (٤٥) وأهمها عدم ملائمة بيانات التقنين ، وعدم كفاءة المعايير (وبخاصة استخدام نسب الذكاء التقليدية) ، وعدم دقة تفسير الدرجات في ضوء المحكات المهنية والتعليمية ، وعدم توافق بيانات كافية عن صدق البطارية وعدم صلاحية الطرق التي استخدمت في حساب ثبات بعض الاختبارات ، فقد استخدمت طريقة التجزئة النصفية في بعض الاختبارات دون مراعاة اعتماد درجاتها اعتمادا كبيرا على السرعة ، ونقصان معاملات ثبات بعض الدرجات العاملية (٢١ : ٢٧٤٠) . ومن المعروف أن أفضل طرق حساب الثبات لهذه الاختبارات التي تعتمد على السرعة طريقتا إعادة الاختبار والضوء المتكافئة ، فإذا اضطر الباحث الى استخدام التجزئة

النصفية فيجب أن تتم التجزئة على أساس السرعة في الأداء لا على أساس الأسئلة في ذاتها وهذا ما أكدته بحث أنستازى ودوريك (٢٢) .

وقد تعرضت بطارية القدرات العقلية الأولية لتعديل آخر عام ١٩٦٢ للتغلب على هذه الصعوبات الفنية الا انه لا تزال أقل في مستواها العلمى من نظائرها من البطاريات ، ناهيك عن تخلفها المعاصر عن التطورات التى طرأت على نظريات التنظيم العقلى وعدد القدرات التى يتألف منها العقل الانسانى . والطبعة الراهنة لهذه الاختبارات تتألف من ٥ بطاريات على النحو الآتى :

- ١ - من مرحلة الحضانة حتى بداية الصف الثانى الابتدائى (من ٥ - ٨ سنوات) .
- ٢ - من الصف الثانى حتى الصف الرابع الابتدائى (من ٨ - ١٠ سنوات) .
- ٣ - من الصف الرابع حتى الصف السادس الابتدائى (من ١٠ - ١٢ سنة) .
- ٤ - تقابل المستويات العمرية للمرحلة الاعدادية فى مصر (من ١٢ - ١٥ سنة) .
- ٥ - تقابل المستويات العمرية للمرحلة الثانوية فى مصر (من ١٥ - ١٨ سنة) .

وتتطلى كل بطارية من هذه البطاريات درجات منفصلة فى القدرات الأولية الآتية :

- ١ - الفهم اللفظى (معانى الكلمات) .
- ٢ - القدرة العددية .
- ٣ - الاستدلال .
- ٤ - السرعة الادراكية .
- ٥ - العلاقات المكانية .

وتتضمن بطارية الأعمار من ١٠-١٢ سنة اختبارات لقياس العوامل الخمس

أما بالنسبة لبطاريات المستوى الأدنى يحذف منها اختبار الاستدلال لصعوبته، وبالنسبة لبطاريات المستوى الأعلى يحذف منها اختبار السرعة الإدراكية لسهولة . وتتحول درجات كل عامل ، وكذلك الدرجة الكلية في البطارية إلى نسب ذكاء انحرافية بدلا من نسبة الذكاء التقليدية في طبعة ١٩٤٧ . وتذكر كراسات التعليمات أن معلمى دور الحضانة والصفوف الأولى من المرحلة الابتدائية يهتمون بالأعمار العقلية أكثر من غيرها من المعايير ولذلك حولت جميع درجات بطارية مرحلة الحضانة حتى بداية الصف الثانى الابتدائى إلى نسب ذكاء تقليدية فقط ، كما حولت درجات بطارية الصف الثانى حتى الصف الرابع الابتدائى إلى نسب ذكاء تقليدية ونسب ذكاء انحرافية . وهذا المطلب واجهته أيضا كما أشرنا مقاييس وكسلر للأطفال في المرحلة الابتدائية WISC وفي درر الحضانة WPPSI ، إلا أن هذه المقاييس أعطت الأعمار العقلية واحتفظت بانحرافات معيارية موحدة لجميع الأعمار، وهذا ما كان يجب على اختبارات القدرات العقلية الأولية عمله بدلا من تغيير وحدات القياس في المستويات العمرية الدنيا . وبالإضافة إلى ذلك تتوافر للبطاريات الأخرى معايير مئينية وتساعية .

وقد أثمرت بطارية القدرات العقلية الأولية ثروة من المعلومات التي توافرت من البحوث الكثيرة التي أجريت عليها في مراحلها الثلاث ، ومع ذلك فإننا نذكر التحفظات الآتية عليها :

١ - أن معاملات الارتباط بين الاختبارات الفرعية التي تتألف منها البطارية مرتفعة نسبيا ، فقد تراوحت هذه المعاملات لبطارية الأعمار من ٨ - ١٠ بين ٤١ ر ، ٧٠ ر ، و لبطارية الأعمار من ٥ - ٨ بين ٥١ ر ، ٧٣ ر . وبالطبع فإن هذه المعاملات المرتفعة لا تتفق مع الأساس الذى قامت عليه بطاريات الاستعدادات المتعددة ، أى الاستقلال النسبى للعوامل التي تقيسها هذه البطاريات . وقد يرجع ذلك إلى عدم استقلال هذه الاستعدادات في هذه المراحل المبكرة من العمر .

٢ - لا تتوافر لهذه البطارية بيانات امبريقية عن صدقها في ضوء المحكات التعليمية والمهنية ، كما لا تتوافر بيانات عن الفروق بين الجنسين في بعض الاختبارات التي تقيس العوامل التي أكدت الدراسات السابقة وجود

فروق بين الجنسين فيها . وبصفة عامة فإن نتائج البحث في صدق هذه البطارية في صورها المبكرة تدل على أن اختبارى معانى الكلمات والاستدلال يرتبطان أعلى من غيرهما من الاختبارات الفرعية بمحكات النجاح المدرسى . ويعلق غريمان (٢٧ : ٤١٥) على ذلك بأن هذين الاختبارين يتشابهان مع ما تضمنته مقاييس الذكاء العام . بل أن اختبارات الذكاء العام - فى رأيه - تتضمن أسئلة تقترب من حيث التعقد والبناء من أنواع النشاط العقلى التى يتطلبها العمل المدرسى أكثر من هذين الاختبارين من بطارية القدرات العقلية الأولية .

٣ - لم تساير بطارية القدرات العقلية الأولية التطورات التى لحقت بنظرية الذكاء فى الوقت الحاضر . ومن ذلك مثلا أن اختبار الاستدلال الذى يستخدم فيها ، وهو من نوع تكلمة سلاسل الحروف ، ليس أفضل اختبارات الاستدلال العام . وكذلك فإن اختبار مراجعة الجمع العددي الذى يستخدم فى هذه البطارية ليس أفضل مقاييس القدرة العددية . بالإضافة الى أن أى بطارية معاصرة للقدرات المتعددة يجب أن تتضمن عدة اختبارات لقياس الأنماط المتعددة للتفكير كما تحددت فى الدراسات العلمية الحديثة (١٩) .

٤ - الخلط الواضح فى بنية هذه البطارية بين قدرات العمليات المعرفية (الاستدلال ، الذاكرة ، الإدراك) وقدرات المحتوى (القدرة اللغوية ، القدرة المكانية ، القدرة العددية) وهو تمييز ليس له ما يبرره فى النظرية الحديثة للقدرات العقلية (١٩) .

وقد أعد هذه البطارية باللغة العربية أحمد زكى صالح وعدلها بما يتفق مع البيئة المصرية ، وتآلف من ٤ اختبارات فرعية هى :

١ - معانى الكلمات : ويتطلب من المفحوص التعرف على الكلمة التى ترادف فى المعنى كلمة معطاة .

٢ - الإدراك المكانى : ويتطلب من المفحوص التعرف على الأشكال المنحرفة والمعكوسة أو المقلوبة .

٣ - التفكير (أو الاستدلال) : ويتطلب من المفحوص تكملة سلاسل حروف .

٤ - العدد : ويتطلب من المفحوص مراجعة صحة عمليات جمع معطاة .

وتعطى للمفحوص درجة في كل اختبار من هذه الاختبارات الأربع ، بالإضافة الى درجة خاصة بالنتائج العام تحسب بالمعادلة الآتية :

$$ق \cdot ع = ا ل + \frac{١}{٢} ك + ا ف + ا ع \cdot$$

حيث الرمز ق · ع = القدرة العامة .

- ا ل = درجة المفحوص في اختبار معانى الكلمات .
- $\frac{١}{٢} ك$ = نصف درجة المفحوص في اختبار الادراك المكاني .
- ا ف = درجة المفحوص في اختبار التفكير .
- ا ع = درجة المفحوص في اختبار العدد .

وتحول الدرجات الخمسة الى تخطيط نفسى (برو فيل) باستخدام معايير البطارية . وقد طبقت البطارية لأكثر من ١٥ عاما على مجموعات كبيرة من الأفراد ، لا يقل عدد الأفراد الذين طبق عليهم الاختبار فى كل سن عن خمسمائة فرد من الجنسين هم العينة الكلية التى استخرجت منها معايير الاختبار ، وفى بعض الأعمار زاد عن ذلك مثل سن السابعة عشرة ، (٩ : ١٠) . واستخرجت المعايير لكل عمر زمنى على حدة من سن ١٣ - ١٧ سنة فى صورة مئينيات ، ونسب ذكاء تقليدية .

وتبين كراسة تعليمات البطارية معاملات ثبات الاختبارات الفرعية وهى تتراوح بين ٨١ر ، ٩٥ر . الا أنها لم تبين حجم العينات التى حسبت منها هذه المعاملات ولا الطريقة التى استخدمت فى ذلك . كما حسبت معاملات الارتباط بين الاختبارات الفرعية ، وهى بصفة عامة معاملات منخفضة مما يدل على الاستقلال النسبى للعوامل . أما عن صدق الاختبارات فتذكر

كراصة التعليمات أنه حسبت معاملات الارتباط بين الاختبارات الفرعية وغيرها من الاختبارات بالإضافة الى التكوين العاملى للاختبارات الفرعية ، ونحن ننصح المهتم بهذه البطارية بالرجوع الى البحوث التى اشرف عليها احمد زكى صالح حتى وفاته عام ١٩٧٤ والتي قام بها طلاب الدراسات العليا بقسم علم النفس بكلية التربية جامعة عين شمس ، للحصول على مزيد من التفاصيل . كما ننصح اى مستخدم لها فى الوقت الحاضر ان يعيد تقنين البطارية حيث أن معظم بياناتها السيكومترية تحتاج الى تحديث .

اختبارات الاستعدادات الفارقة :

ظهرت هذه البطارية لأول مرة عام ١٩٤٧ باسم اختبارات الاستعدادات الفارقة Differential Aptitude Tests (DAT) واعدتها بينيت C.K. Bennett وسيشور H.B. Seashore ووزمان A.C. Wesman ، ثم عدلت عام ١٩٦٢ وأعيد تقنينها ، وظهرت الطبعة الرابعة منها عام ١٩٦٦ .

وتعد هذه البطارية من أكثر بطاريات الاستعدادات المتعددة طموحا ، وتتوافر عنها - فى الخارج - بيانات احصائية من خلال التقنين والتحليل فى دراسات عديدة كثيرة . ورغم أنها تصلح خاصة للمستويات العمرية من ١٤ سنة الى ١٨ سنة ، وفى أغراض التوجيه التربوى والمهنى ، الا انها يمكن ان تستخدم مع الجماعات غير المنتقاة من الراشدين .

وتتألف البطارية من ٨ اختبارات قننت جميعا على نفس عينة التقنين، ومعنى ذلك أن معايير الاختبارات لها نفس الدلالة النسبية . كما أمكن التحكم فى مدى العمر الزمنى والاستعداد والمستوى الدراسى والعوامل غير العقلية فى الشخصية : ومعنى ذلك أن التخطيط النفسى له معناه عند تفسير الفروق داخل الفرد . وتتوافر بطاريات منفصلة ومعايير مستقلة للذكور والاناث فى صورة مئينيات وتساعيات .

وتقاس القدرات المتضمنة فى هذه البطارية قياسا مستقلا ، بالإضافة الى ان الاختبارات التى تتضمنها تعتمد على القوة فى الأداء وليس السرعة، فيما عدا اختبار السرعة والدقة الكتابية .

ورغم أن مؤلفى البطارية لم يستخدموا منهج التحليل العاملى فى اعدادها ، الا أنهم استعانوا فى الاختبارات التى تتألف منها بنتائج البحوث العاملية بالاضافة الى المطالب العملية ، ولذلك فإنهم لم يسعوا الى اختصار الاختبارات « ذات التكوين العاملى البسيط » أو « النقية عامليا » ، وإنما اختاروا الاختبارات المركبة عامليا ما دامت تفيد فى مجالات تربوية أو مهنية لها أهميتها . وتتألف البطارية من الاختبارات الآتية :

١ - الاستدلال اللفظى : وتتألف أسئلته من القياس التمثيلى اللفظى ويتشبع بالمفاهيم والعلاقات اللفظية المركبة ، ومع ذلك فيغلب عليه طابع الاستدلال وليس الفهم اللفظى . لأن الاختبار لا يؤكد صعوبة المفردات اللفظية ، ومن أمثلة أسئلته :

••• بالنسبة الى المساء مثل الفطور بالنسبة الى ••

- « ١ ، العشاء - الركن .
- « ب ، الهادىء - الصباح .
- « ح ، الباب - الركن .
- « د ، الجرى - الاستمتاع .
- « هـ ، العشاء - الصباح .

ويرى فريمان (٢٧ : ٤١٧) أن هذا الاختبار يقيس فى الواقع جانباً محدوداً من الاستدلال اللفظى ، ولذلك يحسن أن يسمى اختبار « التعاثل اللفظى » .

٢ - القدرة العددية : ويشمل هذا الاختبار مدى واسعا من العمليات الحسابية ، ولا يلعب فيه حل المسائل الحسابية أى دور . وتقيس بعض الأسئلة المهارة فى اجراء العمليات الأساسية الأربعة . ويتطلب البعض الآخر فهم المفاهيم والعلاقات الكمية .

٣ - الاستدلال المجرد : ويقىس القدرة على الاستدلال باستخدام المواد غير اللفظية ، ويتألف من سلاسل من الرسوم ، يتطلب كل منها فهم المبدأ الذى يؤدى الى التغير فى الأشكال التى تتألف منها السلسلة أى من نوع سلاسل الأشكال .

٤ - العلاقات المكانية : ويقاس القدرة على التصور البصرى المكانى وذلك بعرض شكلين هندسيين من الأشكال الثنائية البعد ، ويقوم المفحوص بمعالجتهما ذهنيا بحيث يؤديان إلى شكل ثلاثى البعد ، والغرض من الاختبار قياس القدرة على التصور البصرى للأشكال المركبة بعد تدويرها .

٥ - الاستدلال الميكانيكى : ويقاس القدرة على الفهم الميكانيكى بواسطة مراقف مصورة تمثل المبادئ الميكانيكية والعملية ، ويصاحب كل صورة سؤال بسيط موجز حول المبدأ المتضمن .

٦ - السرعة والدقة الكتابية : ويقاس السرعة والدقة فى الاستجابة للاقتراحات الثنائية بين الحروف والأعداد ، ويتطلب المزاجية بين مختلف الاقتراحات مؤكدا ادراك التفاصيل ومعدل الاستجابة . وتهدف الأسئلة التى يتكون منها الى ، الاقتراب من العناصر التى تتألف منها كثير من المهن الكتابية ، .

٧ - الاستخدام اللغوى (القسم الأول) : ويقاس القدرة على تهجى الكلمات، وذلك عن طريق تحديد ما اذا كانت الكلمات التى يتألف منها الاختبار صحيحة أم خاطئة من الناحية الاملائية .

٨ - الاستخدام اللغوى (القسم الثانى) : ويقاس القدرة على استخدام القواعد اللغوية ، وذلك عن طريق تحديد ما اذا كانت الجمل التى يتألف منها الاختبار صحيحة أو خاطئة من ناحية النحو وعلامات الترقيم واستخدام الكلمات .

وتتكون هذه البطارية من صورتين متكافئتين هي الصورة (ل) والصورة (م) . وتعطى كراسة التعليمات الخاصة بالطبعة الرابعة التى ظهرت عام ١٩٦٦ تفاصيل كافية عن الاجراءات التى استخدمت فى بناء البطارية . وقد قننت فى الولايات المتحدة على ٥٠٠٠٠ تليذا فى ١٩٥ مدرسة فى ٤٢ ولاية .

ويذكر السيد محمد خيرى (١٢ : ٢٢١) انه قام بترجمة هذه البطارية

الى اللغة العربية بالاشتراك مع لويس كامل مليكة . وتستخدم في مصر اجزاء مختلفة كثيرة من هذه البطارية لأغراض متعددة منها اختبار الاستدلال الميكانيكى الذى قامت مراقبة الاختبارات السيكولوجية بمصلحة الكفاية الانتاجية باقتباس جزء منه وتعديله بما يتناسب مع أهداف الانتقاء لنظام القلمذة الصناعية فى حرف مختلفة . واختبار الاستدلال اللفظى والسرعة والدقة فى الأعمال الكتابية اللذان اقتبسا وعدلا عند استخدامهما فى اختبارات ديوان الموظفين الذى كان موجودا فى مصر خلال الخمسينات والستينات .

والواقع أنه لا تتوافر لنا أى بيانات تقنين عن هذه البطارية ككل فى اللغة العربية فيما عدا اختبار الاستدلال الميكانيكى ، فقد حسبت المئينيات والدرجات التائية لعينة حجمها ٢٢٢٩ تلميذا فى مراكز التدريب المهنى التابعة لمصلحة الكفاية الانتاجية والتدريب المهنى لوزارة الصناعة (١٢ : ٤٩) .

كما حسب له معامل ثبات بطريقة التجزئة النصفية (١٢ : ١٥) فبلغ ٠٦١ . واستخدم فى حساب صدقه - كجزء من بطارية اختبارات استعدادات حرف المعادن التى سنشير اليها فيما بعد - منهج التحليل العاملى فبلغ مقدار تشبعه « العامل الميكانيكى العام » ٠٥٢ ، ولم يحسب معامل ارتباطه بمعك من محكات التدريب المهنى (١٢ : ٢١) .

بطارية الاستعدادات العامة :

ظهرت بطارية الاستعدادات (General Aptitude Test Battery) (GATB) عام ١٩٤٧ وأعدّها خبراء مكتب التوظيف الأمريكى لأغراض التوجيه والإرشاد فى ميدان الخدمة المدنية ، وتعتمد على افتراض أن « عددا كبيرا من الاختبارات يمكن أن يتجمع فى عدد من العوامل وأن عددا كبيرا متنوعا من المهن يمكن أن يتجمع فى فئات مهنية تبعا لدرجة التشابه بينها فى القدرات المطلوبة » . وييسر هذا اختبار جميع القدرات المهنية عند الشخص فى جلسة واحدة وتفسير درجاته فى ضوء مدى واسع من المهن ، (٢٧ : ٤٢٣) .

وتتألف البطارية من ١٢ اختبارا تعطى ٩ درجات عاملية ، وتتطلب

حوالى ساعتين ونصف فى الاجراء ، والعوامل التسع التى تحدت بمنهج التحليل العاىلى هى :

- ١ - الذكاء العام : ويقاس بالدرجة المركبة من ٢ اختبارات هى :
الادراك المكانى الثلاثى ، والمفردات اللفوية ، والاستدلال الحسابى .
- ٢ - الاستعداد اللفظى : ويقاس باختبار المفردات الذى يتطلب من
المفحوص التعرف على معانى الكلمات (مترادفات) واكدادها .
- ٣ - الاستعداد الحسابى : ويشمل أسئلة تتطلب اجراء العمليات
الحسابية وحل المسائل الحسابية من النوع المعتاد .
- ٤ - الادراك المكانى : ويقاس القدرة على التصور المكانى للأشياء ذات
الأبعاد الثلاثة حين تعرض عرضا ثنائيا .
- ٥ - ادراك الشكل : ويقاس القدرة على المزاوجة بين رسوم آلات معينة
من ناحية وأشكال هندسية من ناحية أخرى .
- ٦ - الادراك الكتابى : ويتطلب المزاوجة بين الأسماء .
- ٧ - التأزر الحركى : ويقاس القدرة على التأزر ، وذلك بأن يطلب من
المفحوص وضع علامات معينة بالقلم الرصاص فى سلسلة من المربعات .
- ٨ - مهارة الأصبع : ويقاس كفاءة المفحوص فى تجميع الأشياء
الصغيرة وفكها .
- ٩ - المهارة اليدوية : ويقاس المهارة اليدوية باستخدام اليدين معا ، ثم
اليد المفضلة ، فى وضع العصى الخشبية فى لوحة العصى ، ثم قلب اوضاع
العصى .

وتتطلب الاختبارات الأربع التى تقاس عاملى مهارة الاصبع والمهارة

اليديوية استخدام بعض الأجهزة البسيطة ، أما الاختبارات الثماني الأخرى فهي من نوع الوراقة والقلم ، وتوجد حور متكافئة للاختبارات السبع الأرنى التى تستخدم فى قياس العوامل الست من الذكاء العام حتى الادراك المكاني .

وتتحول الدرجات العاملة التسع فى بطارية الاستعدادات العامة الى درجات معيارية معدلة متوسطها ١٠٠ وانحرافها المعيارى ٢٠ ، واشتقت هذه المعايير الأمريكية من عينة حجمها ٤٠٠٠ مفحوصا تمثل المجتمع الإحصائى السكانى العام من حيث العمر والجنس والمستوى التعليمى والتوزيع للمهنى والجغرافى ، وباختبار جماعات من الموظفين والمتقدمين للموظائف والمتدربين فى برامج للتدريب لمختلف المهن . وقد أمكن تحديد أنماط الدرجات التى تبين الاستعدادات الفارقة ، والحد الأدنى المطلوب من الدرجة المعيارية لكل مهنة من المهن ، فمثلا وجد أن مهنة المحاسبة تتطلب درجة ذكاء مقدارها ١٠٥ ، ودرجة استعداد عددي مقدارها ١١٥ . أما مهنة السمكرة فتتطلب أن يكون الحد الأدنى من الذكاء ٨٥ ، ودرجة مقدارها ٨٠ فى الاستعداد العددي والاستعداد المكاني والمهارة اليدوية . وهكذا يمكن المقارنة بين بروفيل الدرجات المعيارية للمفحوص فى جميع المهن التى يصل الى درجات للقطع عندها أو يتفوق عليها ، ويستفاد بهذه المقارنة فى أغراض التوجيه والإرشاد المهنى .

وقد قامت وحدة بحوث الدراسات النفسية بالمركز القومى للبحوث الاجتماعية والجنائية (٢٨) ينقل هذه البطارية الى اللغة العربية ، وتم اختيار ٧ من اختبارات البطارية الاثنى عشر لتقنينها على عينات مصرية هى : اختبار البعد (العمليات الحسابية) ، واختبار ادراك العلاقات المكانية ثلاثية البعد ، والمفردات ، واختبار المزاوجة بين الأدوات ، واختبار الاستدلال الحسابى ، واختبار المزاوجة بين الأشكال . واختبار رسم العلامات . وطبقت على عينة من تلاميذ السنة الثالثة الإعدادية ، وتلاميذ المدارس الثانوية العامة والثانوية الفنية ، وتلاميذ مراكز التدريب المهنى التابعة لوزارة الصناعة بلغ عددها ٤٩٨ تلميذا ، واستخدمت درجات هؤلاء التلاميذ فى تحليل الأسئلة التى تتألف منها الاختبارات السبع فحسبت معاملات التمييز ، ودرجات الضعوبة ،

ثم طبقت البطارية في صورتها المعدلة على عينة تقنين كبيرة نسبيا بالمقارنة
ببحوث التقنين التي أجريت على الاختبارات المصرية من النصف الثالث
الاعدادي والصف الثالث الثانوي (علمي وأدبي) ، والتعليم التجارى
والصناعي ، والزراعي ، ومراكز التدريب المهني التابعة لوزارة الصناعة .

وحسبت معاملات الثبات بطريقة اعادة الاختبار فتراوحت بين ٧٩ر،
٩٢ر بالنسبة لمختلف الاختبارات الفرعية داخل نطاق أنواع التعليم المختلفة،
كما استخدمت طريقة التجزئة النصفية لاختبارات القوة فتراوحت معاملات
الثبات بين ٨١ر ، ٩٢ر ، وطريقة كيودر - ريتشاردسون لاختبارات السرعة
فتراوحت معاملات ثباتها بين ٧٥ر ، ٩٦ر . وحسبت المئينيات للعينات الفرعية
التي تتألف منها عينة التقنين الكلية .

وقد حسب صدق التكوين الفرضي عن طريق التحقق من قدرة اختبارات
البطارية على التمييز بين أداء تلاميذ الفئات المختلفة من التعليم . واستخدمت
لهذا الغرض مصفوفة متسلسلة من النسب المخرجة لتحديد قيمة الدلالة
الاحصائية للفروق بين المتوسطات . كما حسب الصدق التنبؤي على عينة من
٢٠٠ تلميذ من مراكز التدريب المهني بأنواعها المختلفة وقورنت درجاتهم في
البرامج التدريبية التي تعرضوا لها قبل تطبيق الاختبارات عليهم وبعده .
وعرضت النتائج في صورة جداول توقع لكل اختبار على حدة ، ومعنى ذلك
أن درجات هذا الاختبار يمكن أن تفسر في ضوء كل من المعيار والمحك .

اختبارات فلاناجان لتصنيف الاستعدادات :

تمثل هذه البطارية اتجاها مختلفا في بناء بطاريات الاستعدادات
المتعددة ، وتسمى (**Flanagan Aptitude Classification Tests (FACT**) .

وقد ظهرت نتيجة لبحوث فلاناجان حول اعداد اختبارات تصنيف افراد
القوات الجوية بالجيش الأمريكي اثناء الحرب العالمية الثانية ، ولذلك فانها
تتجه وجهة عملية تطبيقية ، وتهتم بمجال التوجيه المهني : وقد اعتمدت على
منهج تحليل العمل لكثير من المهن الذي أدى الى تحديد ٢١ عتصرا مهنيا

حاسما ، أو قدرة تميز بين الناجحين والفاشلين من العاملين في كل مهنة .
ويتميز أي عنصر مهني أو قدرة بأنه على قدر من العمومية ، أي يشترك في
أنماط عديدة من المهن ، وفي نفس الوقت على قدر من النوعية بمعنى أنه يقيس
شيئا مختلفا عن العناصر المهنية الأخرى .

وتتألف البطارية من ٢١ اختبارا لقياس العناصر المهنية منها ١٩ اختبارا
كتابيا ، واختباران عمليان هما النقش Carving والنقر Tapping .
وهذه الاختبارات تمثل بطارية غير متجانسة تتضمن وظائف سلوكية مختلفة،
منها الإدراك الحسي والنشاط الحركي والذاكرة والنشاط اللفوي والعدد
والفردات اللفوية العامة والاستدلال اللفوي والاستخدام اللفوي . وفيما
يلي عرض موجز لهذه العناصر التسعة عشر التي تقيسها الاختبارات
اللفظية .

١ - الفحص inspection : وقياس القدرة على تحديد الأخطاء
بسرعة ودقة في سلسلة من رسوم الأشياء (وهو اختبار للإدراك البصري
للتفاصيل) .

٢ - التجميع assembly : وقياس القدرة على التصور البصري
لشيء معين حين توضع مجموعة من الأجزاء معا (وهو اختبار لإدراك
العلاقات المكانية الثلاثية البعد) .

٣ - الحكم والفهم judgment & comprehension : وقياس
القدرة على القراءة مع الفهم واستخدام الحكم السليم في المواقف العملية
(اختبار لفهم معاني الفقرات) .

٤ - الحنق ingenuity : وقياس القدرة على الابتكار والاختراع
في إعداد إجراءات أو آلات أو عروض ذكية (اختبار استدلال يعتمد على
مشكلة محددة) .

٥ - اليقظة alertness : وقياس القدرة على تقدير الموقف وملاحظة
ما ينشأ فيه من أخطار أو ما يحتاجه من وسائل وخطط (اختبار لإدراك
التفاصيل في صورة معينة وما بينها من علاقات) .

٦ - الترميز coding : ويتطلب السرعة والدقة فى ترجمة معلومات مكتوبة الى لغة شفوية معينة .

٧ - الذاكرة memory : ويقيس القدرة على استدعاء الرموز المتعلمة من اختبار الترميز السابق .

٨ - الدقة precision : ويتطلب السرعة والدقة فى اصدار حركات دقيقة للاصبع .

٩ - المقاييس Scales : ويتطلب السرعة والدقة فى قراءة المقاييس والرسوم البيانية .

١٠ - التآزر Coordination : ويقيس القدرة على التآزر بين حركات اليد والذراع .

١١ - الحساب arithmetic : ويتطلب الكفاءة فى العمليات الحسابية الأساسية .

١٢ - الأنماط patterns : ويقيس القدرة على محاكاة واسترجاع الخطوط الرئيسية لنمط بسيط مرسوم .

١٣ - المكونات components : ويقيس القدرة على تحديد الأجزاء الهامة فى مجموعة من الرسوم والتخطيطات .

١٤ - الجداول tables : ويقيس القدرة على قراءة نوعين من الجداول أولهما يستخدم الأعداد ، والآخر يستخدم الكلمات والحروف الأبجدية .

١٥ - الميكانيكا mechanics : ويقيس القدرة على فهم المبادئ الميكانيكية .

١٦ - التعبير expression : ويقيس القدرة على نقل الأفكار كتابة او شفويا .

١٧ - الاستدلال reasoning : وقيس القدرة على حل المشكلات والتعبير عن الحل في صورة رياضية .

١٨ - المفردات vocabulary : وقيس القدرة على فهم معاني الكلمات .

١٩ - التخطيط planning : ويتطلب الكفاءة في تنظيم البيانات وذلك باعادة ترتيب خطوات معينة لحل المشكلة .

ورغم أن الاختبارات التي تتألف منها البطارية (وعددها ٢١ كما ذكرنا) تصحح منفصلة ، فإن درجات مختلف الاختبارات تجمع بعضها الى بعض عند التنبؤ بالنجاح في مهن معينة ، وتتفاوت هذه المهن من المستوى المهني الرفيع (كالطب والهندسة والتدريس) الى مستوى العامل الماهر . وحيث أن هذه الاختبارات المنفصلة تمثل تنوعا واختلافا في العمليات النفسية التي تقيسها ، بالإضافة الى انخفاض معاملات الارتباط فيما بينها ، فإن جمع درجات هذه الاختبارات يصبح ممارسة تثير الكثير من التساؤل والشك .

وتتطلب بطاريات الاختبار اللفظية زمنا كليا طوله عشر ساعات ونصف تقسم الى ٣ جلسات اختبارية . وقد أعدت معاييرها الأمريكية على عينة كبيرة من طلاب المرحلة الثانوية ، ومع ذلك فإنها يعوزها التمثيل كما تقرأ انستازي (٣١ : ٣٤٢) .

وبرى فلاناجان أن منهج العناصر المهنية يقع في منزلة متوسطة بين تحليل القدرات الى عوامل اولية (او نقية) من ناحية وطريقه عينة العمل التي تتم فيها المماثلة بالعناصر الأساسية في الأعمال والمهن الحقيقية من ناحية أخرى .

وتؤكد نتائج البحوث التي قام بها فلاناجان (٣٥) أن معاملات الارتباط بين الاختبارات الفرعية منخفضة بوجه عام حيث يبلغ وسيطها ٢٠ في عينة من الطلاب من سن ١٥ سنة ، و ٣١ في عينة من الطلاب من سن ١٨ سنة . ومعنى هذا أن الصدق الداخلي للبطارية قد تحقق (أى الاستقلال النسبي للاختبارات الفرعية) .

وتؤكد نتائج الصدق التلازمى باستخدام معك التحصيل المدرسى عدم اللجوء الى استخدام الاختبارات الفرعية بسبب انخفاض معاملات الارتباط بصفة عامة ، كل على حدة ، فى اغراض الترجييه التعليمى . والواقع أن الدرجة المجمعة لدرجات جميع الاختبارات أدت الى معاملات ارتباط متعددة لها دلالة احصائية . الا ان الصعوبة الجوهرية فى اللجوء الى البطارية ككل تكمن فى الوقت الطويل الذى تحتاجه فى تطبيقها .

وقد قام فلانجان بدراسات تتبعية لأفراد عينة من طلاب الجامعة فى عدد من مجالات الدراسة ، بعد أن اختبرهم فى أثناء دراستهم الثانوية (لحساب الصدق التنبؤى) ، وبلغ وسيط معاملات الارتباط ٢٩ر٠ . وكانت أقل معاملات الارتباط لطلاب الدين والخدمة الاجتماعية وأعلها لطلاب العلوم الاجتماعية (التى شملت ٦ منها على سبيل المثال الاختصاص النفسى ، ورجل القانون ، والمؤرخ) . كما حسبت معاملات الارتباط بين درجات الاختبارات الفرعية وتقديرات الرؤساء والمشرفين فى عدد من المهن ، فوجد أن هذه المعاملات منخفضة بصفة عامة ، بل قد تصل فى بعض الأحيان الى الصفر أو عدم الدلالة الاحصائية .

وقد أجريت على هذه البطارية دراستان فى مصر أحدهما على طلاب كلية الهندسة (١١) والأخرى على طلاب كلية الطب (١٢٧) .

بطارية اختبارات استعدادات حرف المعادن :

أعدت بطارية اختبارات استعدادات حرف المعادن فى مصلحة الكفاية الانتاجية والتدريب المهنى بوزارة الصناعة بإشراف السيد محمد خيرى (١٢) . وتطلب إعدادها تحليل الأعمال التى يدرّب عليها تلاميذ مراكز التدريب المهنى التابعة لهذه المصلحة ، وأسفرت نتائج هذا التحليل بالنسبة لجميع حرف المعادن والسيارات والتبريد على استخلاص القدرات المهنية التالية : العلاقات المكانية ، الفهم الميكانيكى ، التصور البصرى ، تذكر الأشكال ، سرعة الإدراك ، المعلومات أو الخبرة الميكانيكية ، تقدير الأطوال والمساحات والحجوم ، ثبات اليد ، مهارة الأصابع . وأعدت بطارية اختبارات لقياس هذه القدرات تتألف من قسمين أحدهما يعتمد على ورقة وقلم والآخر من النوع العملى ، وفيما يلى وصف موجز لهذه الاختبارات .

اختبارات الورقة والقلم : وتشمل ٨ اختبارات هي :

١ - الاستدلال اللفظي : وقيس القدرة على الاستدلال اللفظي وإدراك المفاهيم اللفظية والقدرة على التجريد ويتألف من ٣٠ سؤالاً ويستغرق ١٥ دقيقة . ويتكون كل سؤال من جملة تتضمن مفهوماً معيناً أو علاقة معينة على المفحوص اكتشافها بالاختيار من متعدد .

٢ - اختبار الذكاء الاعدادي : للسيد محمد خيرى .

٣ - اختبار الاستدلال الميكانيكى : وهو مقتبس من بطارية اختبارات الاستعدادات الفارقة التى أشرنا إليها آنفاً .

٤ - اختبار المعلومات الميكانيكية . ويهدف الى معرفة مدى الفة المفحوص بالعدد والآلات الميكانيكية ويتكون من ٥ أسئلة فى كل منها ست صور لعدد ميكانيكية بسيطة ، ويطلب من المفحوص التعرف على اسمائها بالاختيار من متعدد .

٥ - اختبار المعلومات الحسابية : وقيس القدرة على التعامل مع الأعداد وذلك بإجراء العمليات الحسابية الأساسية (الجمع والطرح والضرب والقسمة باستخدام الأعداد الصحيحة والكسور الاعتيادية) .

٦ - اختبار التصور المكاني : وهو مقتبس من بطارية الاستعدادات العامة التى أشرنا إليها آنفاً .

٧ - اختبار تكميل الأشكال : وهو مقتبس من بطارية المعهد القومى البريطانى لعلم النفس الصناعى وقيس القدرة على إدراك العلاقات الميكانيكية الثنائية البعد ، والقدرة على المزاوجة السريعة بين الأشكال وإصدار الأحكام عليها بطريقة تعتمد على التفكير الاستنباطى وسرعة الإدراك وتقدير المساحات ، ويتطلب فى الأمثلة الصعبة التصور البصرى .

٨ - اختبار تذكر الأشكال : من اعداد السيد محمد خيرى ، وقيس القدرة على تذكر الأشكال الهندسية بسرعة .

الاختبارات العملية : وتشمل ٣ اختبارات هي :

٩ - اختبار التجميع الميكانيكى : وهو مقتبس من أصل المانى ، ويقس القدرة الميكانيكية وخاصة القدرة على ادراك العلاقات والمتعلقات الميكانيكية والتصور المكانى . ويعتمد على اختبارات مينيسيوفا للقدرة الميكانيكية . ويعتمد على تجميع الأشياء .

١٠ - اختبار مهارة الأصابع : وهو من تصميم محلى يجمع بين جهاز أو أكثر لمهارة الأصابع اختيرت من عدد من الاختبارات الموجودة بالفعل ، ويقس القدرة على تحريك الأصابع وتناول الأشياء الدقيقة بسرعة ودقة .

١١ - اختبار ثبات اليد : ويقس القدرة على التحكم فى حركة اليد .

وقد قننت هذه البطارية على عينة كبيرة تتراوح أعمارها بين ١٤ ، ١٨ سنة من الحاصلين على الشهادة الإعدادية ، وكان تعدادها ، ينحصر بين ألف وثلاثة آلاف تلميذ على حسب الظروف التى طبقت فيها الاختبارات المختلفة ، (٣ : ١٠) ، وشملت قلاميذ المدارس الإعدادية والثانوية والمتقدمين لمراكز التدريب المهنى وتلاميذ هذه المراكز . وحللت أسئلة الاختبارات وتحددت لها معاملات التمييز ، ورتبت حسب مستويات الصعوبة .

وقد حسب الثبات لمعظم الاختبارات بطريقة للتجزئة النصفية ، أما بالنسبة للاختبارات التى تعتمد على السرعة وهما اختبارا مهارة الأصابع وثبات اليد فقد استخدمت معها طريقة إعادة الاختبار . مع ملاحظة أن اختبار العمليات الحسابية لم يبين معامل ثباته فى التقرير المنشور . وكانت معاملات الثبات مرضية بصفة عامة .

ولحساب صدق التكوين الفرضى استخدم منهج التحليل العاظمى لاختبارات الورقة والقلم فقط واستخرج عامل عام يكشف عن مدى تشبع كل اختبار بالاستعداد الميكانيكى العام وكانت التشبعات على النحو الآتى :

الاستدلال اللفظي ٥٧ر ، والذكاء الاعدادي ٥٨ر ، الاستدلال الميكانيكي ٥٢ر ، المعلومات الميكانيكية ٢٧ر ، والعمليات الحسابية ٢٧ر ، التصور المكاني ٥٨ر ، تكميل الأشكال ٥١ر تذكر الأشكال ٤٧ر .

وتم حساب الصئق التلازمى بين الدرجة الكلية فى البطارية ومحتات التحصيل والكفاءة فى مراكز التدريب المهنى فبلغ معامل الارتباط بينها وبين درجات المواد النظرية ٢١ر . وبين درجات المواد العملية ١٨ر وكانت معاملات الارتباط بين اختبارات الورقة والقلم والمواد النظرية ١٩ر ، والارتباط بين الاختبارات العملية والمواد العملية ٤٠ر

واستخرجت معايير الاختبارات الفرعية التى تتألف منها هذه البطارية من اداء عينات التقنين حيث استخرج لكل اختبار معاييره المثينة ودرجاته القائية . ومما يجدر ذكره أن هذه البطارية لم يجز لها أى تقنين جديد منذ ذلك الحين ، ومع ذلك لا تزال تستخدم حتى الآن فى مصلحة الكفاية الانتاجية والتدريب المهنى ، وهو أمر يشكك فى جدواها فى اتخاذ قرارات الانتقاء بعد مرور ما يقرب من ١٥ عاما على آخر دراسة منظمة لها .

بطارية الاختبارات الفارقة لمورسبى :

أعد هذه البطارية فى الأصل العالم الانجليزى ج . و . مورسبى وأصدرتها هيئة البحوث التربوية فى انجلترا وويلز (NFER) وتتألف من الاختبارات الفرعية الآتية :

١ - اختبار المسلسلات المركبة : وفكرته اشبه باختبارات المصفوفات المتتابعة لرافن .

٢ - اختبارات القدرة العامة اللفظية والعديدية والادراكية .

٣ - اختبارات السرعة وتتألف من ٦ اختبارات لقياس القدرة الكتابية والمرونة العقلية والطلاقة اللفظية والطلاقة الفكرية والسرعة اليدوية والمهارة اليدوية .

٤ - اختبار الأشكال .

٥ - اختبار القدرة الميكانيكية

وقد نقل هذه الاختبارات الى اللغة العربية وقتننها على البيئة المصرية مصطفى الشرقاوى (٢٨) كما قنن هذه الاختبارات أيضا على البيئة الجزائرية والليبية .

وكانت العينات المصرية تتألف من طلبة مراكز التدريب المهني بالقاهرة، اعد عليهم تطبيق البطارية وبلغت معاملات الثبات لاختبار المسلسلات المركبة ٨٥ر ، والقدرة العامة اللفظية ٦٧ر ، والقدرة العامة العددية ٧٥ر ، والقدرة العامة الادراكية ٧١ر . اما اختبارات السرعة فتراوحت معاملات ثباتها بين ٦٩ر للطلاقة الفكرية ، ٩٧ر للطلاقة اللفظية ، واختبار الأشكال ٨١ر ، واختبار القدرة الميكانيكية ٩٣ر .

وحسب الصدق للاختبارات البطارية باستخدام طريقة المقارنة الطرفية وتبين ان الاختبارات تميز بين التلاميذ الممتازين والضعاف في دراستهم بمراكز التدريب المهني في بعض الحرف . وحسبت المعايير من أداء عينة تتألف من طلاب المدارس الثانوية العامة بالقاهرة وتلاميذ دبلوم مراكز التدريب المهني بالقاهرة وطلاب السنة الأولى بكلية الاداب والعلوم الاجتماعية بجامعة وهران بالجزائر ، وطلاب المدارس الثانوية بمكتب تعليم طرابلس بليبيا تمت اعمارهم بين ١٧ - ١٨ سنة . واستخدم الباحث معيار نسبة الذكاء الانحرافية في معظم الاختبارات .

بطارية الاختبارات المعرفية مرجعية العوامل :

هذه البطارية من اعداد هيئة خدمات الاختبارات التربوية (ETS) في برنستون بالولايات المتحدة الأمريكية وتسمى بالانجليزية Kit of factor-referenced cognitive tests وظهرت منها عدة طبعات اخرها ما صدر عام ١٩٧٦ ، وقد نقل هذه البطارية الى اللغة العربية فريقان من الباحثين : الأول يتكون من سليمان الخضرى وآنور الشرقاوى ونادية عبد

السلام بجامعة عين شمس) والثانى على حسين بدارى وانور رياض عبد الرحيم (بجامعة المنيا) ولم يتعد الأمر فى حدود علم مؤلف هذا الفصل - حدود الترجمة أو التعريب ، ولا تتوافر لدينا أى معلومات سيكومترية عن البطارية أو مكوناتها .

تقويم بطاريات الاستعدادات المتعددة :

من الواضح ان بطاريات الاستعدادات المتعددة - التى تناولنا نماذج منها فى هذا الفصل - تختلف فى الأسس التى تستند إليها عن اختبارات الذكاء العام . فاختبارات الذكاء العام تتضمن عينات من النشاط العقلى التى لها أهميتها فى المواقف التى تتلاءم معها هذه الاختبارات ، أما اختبارات الاستعدادات المتعددة فإنها تسعى إلى قياس العمليات العقلية المحدودة النطاق والتى تعد ضرورية فى مقررات دراسية معينة أو فى أنماط معينة من المهن . وقد أدى اهتمام بعض علماء النفس ، بالتكوين العائلى البسيط ، إلى استخدام الاختبارات التى تتكون من أسئلة محدودة القيمة فى قياس النشاط العقلى المعقد والضرورى فى عمليات التعلم المدرسى وفى النشاط المهنى . وللتغلب على هذه الصعوبة لجأ بعض صناع بطاريات الاستعدادات المتعددة إلى جمع درجات الاختبارات الفرعية التى تتألف منها هذه البطاريات بطرق مختلفة ولمختلف الأغراض . وفى هذا نوع من الاعتراف بأن العوامل النوعية أو القدرات الخاصة لا تعمل منفردة ، وإنما تنتظم مع غيرها من القدرات فى نمط من النشاط العقلى على درجة من التعقد النسبى .

وتؤكد نتائج البحوث أنه عند استخدام الاختبارات فى الأغراض التطبيقية - كالتوجيه والانتقاء والتصنيف فى مؤسسة معينة ولمهن خاصة - فإن الاختبارات التى تعد خصيصاً لذلك تتفوق على الاختبارات المقننة والمعدة للاستخدام العام . ويصدق هذا أيضاً على الميدان التربوى ، ووراء هذا الرأى ذلك الاهتمام المتزايد فى الوقت الحاضر بالاختبارات التحصيلية . وكذلك فإن الاختبارات التى تقنن على عينات متنوعة من فئات مهنية وتعليمية مختلفة تفيد أكثر من غيرها فى هذه الأغراض العملية . وهذا ما حدث بالفعل فى تقنين معظم بطاريات الاستعدادات المتعددة .

وحينما يستخدم الأخصائى النفسى بطاريات الاستعدادات المتعددة فى

التمييز بين الأفراد، في مدى صلاحيتهم لأنواع التعليم والمهن المختلفة فعليه أن يتيقن من الصدق الفارق differential validity للبطاريات قبل استخدامها فقد يجد أن بعض الاختبارات الفرعية أقدر على التنبؤ بنتائج معينة من غيرها وهنا يقع صانعوا هذه البطاريات في خطأ فادح حين يؤكدون في تقنياتها الفروق بين متوسطات درجات الاختبارات الفرعية التي تتألف منها بالنسبة للجماعات المهنية والتعليمية المختلفة ، ويهملون ما يمكن أن يوجد بين درجات هذه الجماعات من تداخل ، فبين طلاب الطب والهندسة مثلا على الرغم من اختلافهما كثير من العناصر المشتركة ، ومن هنا يتوافر لهم أنماط متشابهة من اختبارات العوامل المتعددة ، إلا أن بينهم أيضا فروق في السمات العقلية وغير العقلية التي لا تقيسها هذه البطاريات .

الفصل العاشر

اختبارات القدرات المتخصصة

أشرنا فيما سبق إلى أن اختبارات الذكاء العام لا تشمل فيما تقيسه بعض « القدرات العقلية » ، وقد ظهر هذا النقص حتى قبل الاهتمام ببطاريات الاستعدادات المتعددة ، وتمثل ذلك في ظهور الاختبارات التي تقيس ما كان يسمى « الاستعدادات الخاصة » ، وأشهرها وأكثرها تبيكراً اختبارات « الاستعداد الميكانيكي » . ثم تزايدت الحاجة إلى هذه الاختبارات التي تقيس ما يمكن أن نسميه « القدرات المتخصصة Specialized abilities » بزيادة الاهتمام في علم النفس التطبيقي بمشكلات التوجيه والانتقاء والتوزيع والارشاد النفسي . فتكاثر عددها وشملت مجالات أكثر تنوعاً ومن ذلك ميادين النشاط الموسيقي والفني والمهني والكتابي ، بالإضافة إلى النشاط الميكانيكي الذي أشرنا إليه .

وقبل أن نتناول نماذج من اختبارات القدرات المتخصصة بشيء من التفصيل لابد أن نوضح مكانة هذه الفئة من الاختبارات في النظرية الحديثة للقدرات العقلية . وبإيجاز نقول أن هذه الاختبارات ظهرت في وقت كان الاهتمام كله مركّزاً على ما يسمى « اختبارات الذكاء العام » ، بهدف أن تكون مكملة لمفهوم « نسبة الذكاء » ، في وصف النشاط العقلي للمفحوص . إلا أنه مع التقدم في مناهج البحث وفي نظرية القدرات العقلية ، وخاصة بعد شيوع استخدام منهج التحليل العاملي والنماذج المؤسّسة عليه (١٩) . أصبح من الشائع اعتبار مفهوم « الذكاء » ذاته يتألف من عدد كبير من « القدرات العقلية » ، يصل في نموذج جيلفورد (١٩ ، ٢٨) إلى أكثر من ١٢٠ قدرة ويزيد عن ذلك كثيراً في النماذج المعاصرة لتجهيز المعلومات (١٩) وزاد نتيجة لذلك الاهتمام ببطاريات الاستعدادات المتعددة . بل إن بعض الاختبارات التقليدية « للاستعدادات الخاصة » أصبحت من مكونات بعض بطاريات الاستعدادات كما أشرنا في الفصل السابق .

والسؤال الذى طرحه أنستازى (٢١) هو : ما هو موضع هذه الاختبارات فى الوقت الحاضر ؟ الواقع ان لهذه الاختبارات وظيفتين هامتين فى التقويم النفسى المعاصر هما :

١ - توجد بعض مجالات السلوك الانسانى لا تتوافر لها مقاييس جيدة فى بطاريات الاستعدادات المتعددة ومن ذلك الادراك البصرى والسمعى والمهارة الحركية والمواهب الموسيقية والفنية والسلوك الابتكارى . وقد يكون السبب فى ذلك ان المواقف التى تتطلب استخدام هذه الاختبارات على درجة كبيرة من النوعية بحيث لا تقرر ادماجها فى بطاريات الاستعدادات المتعددة . ومعنى ذلك ان هذه الاختبارات منتظلة غنة متميزة من الاختبارات العقلية حتى تتوافر بطاريات شاملة لقدرات الانسان . وهذا ما نتنبأ بحداثته فى المستقبل القريب لعلم القياس النفسى .

٢ - تستخدم اختبارات القدرات المتخصصة حتى فى الأحوال التى يوجد لها ما يقابلها فى بطاريات الاستعدادات المتعددة ، ومن ذلك اختبارات القدرات الميكانيكية أو الكتابية . والسبب فى ذلك ان بعض اختبارات القدرات المتخصصة تتوافر لها بيانات تقنين وافية ، بالإضافة الى المرونة فى استخدام الاختبارات ، وشمول الميدان الذى تقيسه . فبدلاً من استخدام اختبار واحد للقدرة الميكانيكية فى احدى البطاريات المتعددة الاختبارات ، يمكن للباحث ان يطبق بطارية لاختبارات القدرة الميكانيكية اذا تطلب الامر قياس مختلف جوانب هذه القدرة فى المفحوصين .

ويتوافر فى الوقت الحاضر عدد هائل من الاختبارات التى تقيس القدرات المتخصصة ، ابتداء من المستوى الحامى (السمع والبصر) حتى مستوى التفكير الابتكارى والنشاط المهنى . ولا يتسع المقام لتناول هذه الاختبارات باى نوع من العرض الشامل ولذلك سوف نقتصر على طائفة من هذه القدرات وما يقيسها من اختبارات .

اختبار القدرات الحركية :

ظهرت الاختبارات التى تقيس خصائص الاستجابات الحركية كالمسرة

والتأخر منذ فترة طويلة ، ويهتم معظمها بالمهارة اليدوية Manual dexterity وقليل منها يهتم بحركات الساق والقدم والتي تتطلبها بعض المهن والأعمال .
ويقيس بعضها مزيجاً من الاستعدادات الحركية والادراكية والمكانية والميكانيكية . وفي أغلب الأحوال فإن اختبارات القدرات الحركية تتألف أجهزة في قياسها رغم ظهور بعض اختبارات الورقة والقلم في أغراض التقويم النفسى الجماعى . وقد أشرنا فى الفصل الخامس الى بطاريتى فلانجان والاستعدادات العامة اللتين تضمنتا بعض الاختبارات الحركية الجماعية الا ان الدراسات التى قام بها فلشمان (٢٦) وملتون (٤١) تؤكد عدم وجود علاقة بين اختبارات الورقة والقلم واختبارات الأجهزة التى تقيس نفس القدرات الحركية ، وفيما يلى امثلة للأجهزة التى تستخدم فى قياس النشاط الحركى :

١ - مقياس القوة العضلية (الدينامومتر) Dynamometer
وتستخدم فى قياس قوة قبضة اليد .

٢ - جهاز قياس زمن الرجوع .

٣ - اختبار Crawford Small Parts Dexterity و يقيس المهارة
اليدوية البسيطة .

٤ - اختبار Purdue Pegboard و يقيس النشاط الحركى الدقيق
والنشاط الحركى الغليظ .

٥ - اختبار Stromberg Dexterity و يقيس حركات اليد والذراع .

اختبارات القدرات الميكانيكية :

ان ما يسمى القدرة الميكانيكية ليس قدرة موحدة لا تقبل التحليل ، وانما هى فى الواقع مركب من القدرات الحسية والادراكية والحركية ، بالإضافة الى قدرات ادراك العلاقات المكانية ، واكتساب المعلومات الميكانيكية ، وفهم العلاقات الميكانيكية .

ولذلك فان الاختبارات التى تقيس القدرة الميكانيكية انما تهتم بمستوى من الأداء اعلى مما تتضمنه الاختبارات الحسية والادراكية (التمييز السمعى والبصرى والحركى والتأزر العضلى والمهارة اليدوية) .

وبعد اختبار التجميع assembly الذى اعد ستينكويست Stenquist عام ١٩٢٣ اول اختبار ظهر فى هذا الميدان يقيس قدرة المفحوص على تركيب اجزاء الأجهزة الميكانيكية مثل جرس الدراجة ، او قفل ، او مصيدة فئران . ويتألف من ٣ سلاسل تستخدم مع مستوى عمرى يعقد بين الطفولة والرشد .

وقد عدل هذا الاختبار فى جامعة مينيسوتا عام ١٩٣٠ واصبح يطلق عليه اختبار مينيسوتا للتجميع الميكانيكى Minnesota Mechanical Assembly Test واضيفت اليه أجهزة جديدة . ويصحح هذا الاختبار كسابقه فى ضوء معدل الاستجابة ودقتها . وقد اكدت الدراسات التى أجريت عليه انه على درجة كافية من التنبؤ بنجاح طلاب التعليم الفنى المبتدئين فى اعمال الورشة ، كما ارتبط بالنجاح فى بعض المهن الميكانيكية .

وفى عام ١٩٣٠ ظهر اختبار آخر فى جامعة مينيسوتا هو اختبار مينيسوتا للعلاقات المكانية Minnesota Spatial Relations Test ويتألف من سلسلة من ٤ لوحات خشبية يتكون كل منها من ٥٧ قطعة خشبية مختلفة الأشكال ، وبعضها غير مألوف ويطلب من المفحوص أن يضع هذه القطع فى اماكنها فى فتحات اللوحة الخشبية ، وفى رأى كثير من النقاد ان هذا الاختبار يقيس السرعة والدقة فى الاستجابة لتفاصيل العلاقات المكانية . وقدرة المفحوص على التعامل مع الأشياء والمواد المحسوسة . وبعبارة أخرى فانه لا يقيس القدرة على حل المشكلات ذات الطبيعة الميكانيكية . كما لا يصلح لقياس القدرة على معالجة شياء الصغيرة بدقة .

وفى عام ١٩٤٨ ظهر اختبار مينيسوتا للوحات الاشكال الورقية
Revised Minnesota Paper formboard

يعرض على المفحوص على ورقة مطبوعة نفس المشكلات التى كانت تعرض عليه فى لوحات الاشكال الحقيقية . وفى كل سؤال من أسئلة هذا الاختبار يعرض على المفحوص جزءان أو اكثر من شكل هندسى بحيث

لو جمعت معا بطريقة صحيحة فانها تكون الشكل الكلى الكامل . وعلى
المفحوص تحديد الشكل الذى يمكن أن يجمع مع الأجزاء الأخرى من بين ٥
بدائل اختيارية . وهكذا يقيس هذا الاختبار القدرة المكانية كما تحددت فى
البحوث العاملة للقدرة العقلية . وتؤكد نتائج البحوث التى أجريت على
هذا الاختبار أنه يرتبط ارتباطا دالا بنوع الأداء الميكانيكى والنجاح فى الرسم
الهندسى والهندسة الوصفية . واستخدمت فى حساب صدقة محكات مختلفة
مثل درجات مقررات الهندسة والمواد الفنية والورش الميكانيكية وتقديرات
المشرفين ومقاييس الكفاءة الانتاجية . وبوجه عام يمكن القول أن طلاب
الهندسة والعاملين بالمهن الميكانيكية يحصلون على درجات فى هذا الاختبار
أعلى من غيرهم ، وقام بتقنيته على البيئة المصرية فؤاد أبو حطب وأمال
صديق (٢١) .

وتوجد فئة أخرى من اختبارات القدرة الميكانيكية تقيس المعلومات
الميكانيكية والاستدلال الميكانيكى والفهم الميكانيكى . وهذه جميعا تتطلب
من المفحوص الفة بالآلات الشائعة وبالعلاقات الميكانيكية ، إلا أنها لا تفترض
فى المفحوص توافر معلومات فنية أكثر مما يمكن أن توفره الحياة اليومية
فى المجتمعات الصناعية . ومن أشهر هذه الاختبارات تلك التى أعدها بينت
وزملاؤه (٢٢) والتى تسمى اختبارات الفهم الميكانيكى Tests of Mechanical
Comprehension

وتعرض على المفحوص مشكلات ميكانيكية مصورة وعدة حلول لكل
منها ، وعليه أن يختار الحل الصحيح . ولهذه الاختبارات ثلاثة مستويات
من الصعوبة ، وتهدف الى قياس القدرة على فهم المبادئ الفيزيائية
والميكانيكية فى مواقف بسيطة نسبيا .

ومن الاختبارات الأحدث نسبيا فى قياس القدرة الميكانيكية اختبار
الدافعية الميكانيكية المسمى Mellanbruch Mechanical Motivation Tests
وهو اختبار مصور يتطلب من المفحوص التعرف على أشياء معينة شائعة فى
المجتمعات الصناعية المتقدمة . والاختبار الثانى هو SRA Mechanical
Aptitudes Test ويتألف من ٢ اختبارات فرعية تقيس :

(التقويم النفسى)

- (١) المعلومات الميكانيكية .
- (ب) ادراك الاشكال والتصور البصرى المكانى .
- (ح) حل المشكلات الحسابية الخاصة بالأعمال الميكانيكية (فى صورة جداول ورسوم بيانية) .

ويجب ان نؤكد ان هذه الاختبارات التى تقيس الفهم الميكانيكى تتطلب تعديلا شاملا عند محاولة استخدامها فى البيئة المصرية ، وذلك لأنها مشبعة تشبعا كبيرا بالعناصر الثقافية السائدة فى المجتمعات التى صنعت فيها .

اختبارات القدرة الكتابية :

القدرة الكتابية Clerical ability - كالقدرة الميكانيكية - ليست قدرة بسيطة ، وانما هى على درجة من التعقد والتركيب . ورغم أن بعض بطاريات الاستعدادات المتعددة التى تناولناها فى الفصل الخامس - وخاصة بطارية الاستعدادات الفارقة وبطارية الاستعدادات العامة - تضمنت اختبارات فرعية لقياس القدرة الكتابية فان تراث التقويم النفسى يتضمن طائفة من الاختبارات التى تقيس هذه القدرة على وجه الخصوص ، بل توجد بطاريات لقياس القدرة الكتابية . وعموما نستطيع القول أن اختبارات هذه القدرة تجمعها خصائص مشتركة يوضحها الجدول الآتى الذى يلخص محتوى ٦ بطاريات للقدرة الكتابية واختباراتها الفرعية . ومنه يتضح أن البطاريات الشاملة تسعى الى قياس عدد من العمليات النوعية . بينما تسعى البطاريات الأقل شمولاً الى قياس السرعة والدقة الإدراكيّتين والتى تعد أحد جوانب القدرة الكتابية وليست كلها . وقد يكون من العوامل الكامنة وراء قصر القدرة الكتابية على السرعة الإدراكية ما تؤكدّه نتائج البحوث من أهمية القدرات الإدراكية بوجه عام فيها . إلا أن « القدرة الكتابية » تشمل أنواعاً متنوعة من الأعمال تختلف فى مدى القدرة ومستواها . ولذلك فإن مقاييسها تضمن عمليات عقلية أخرى أساسية ، بالإضافة الى بعض القدرات الحسية والمهارات اليدوية عند الضرورة . ويلخص الجدول رقم (١٥) الاختبارات الفرعية التى تتضمنها ٦ بطاريات للقدرة الكتابية .

جدول رقم (١٥) يلخص الاختبارات الفرعية التي تتضمنها ٦ بطاريات للقدرة الكتابية .

البيطارية	الاختبارات الفرعية
بطارية ديترويت Detroit	<p>المخط : السرعة والدقة المراجعة : السرعة والدقة الحساب البسيط السرعة والدقة في الأداء الحركي معرفة المصطلحات التجارية البسيطة ترتيب الصور التصنيف : السرعة والدقة الترتيب الأبجدي</p>
بطارية الاستعدادات الكتابية العامة General Clerical	<p>المزاوجة : اكتشاف الأخطاء في الأسماء والأعداد الترتيب الأبجدي الحساب : الجمع البسيط الحساب : تحديد أخطاء الجمع المسائل الحسابية البسيطة التهجي فهم القراءة معرفة معاني الكلمات الاستخدام اللفوي : القواعد النحوية</p>

تابع الجدول رقم (١٥)

جدول يلخص الاختبارات الفرعية التي تتضمنها ٦ بطاريات للمقدرة الكتابية •

البطارية	الاختبارات الفرعية
بطارية مينيسوتا	مقارنة الأعداد مقارنة الأسماء
بطارية برديو Purdue	التهجى الحساب المراجعة : السرعة معرفة معانى الكلمات النقل والنسخ : الدقة الاستدلال
بطارية التوظيف Short Employment	العمليات العددية معانى الكلمات التصنيف
بطارية تورس Turse	المهارات اللفظية المهارات العددية التعليمات المكتوبة المراجعة : السرعة التصنيف الترتيب الأبجدي

وتوجد فى المكتبة العربية بطاريتان لقياس القدرة الكتابية هما :

(١) اختبارات القدرة الكتابية وهى من اعداد محمد عبد السلام احمد (٢٤) عام ١٩٦٠ وتآلف من اختبارين هما : اختبار تصنيف الأعداد ، واختبار تصنيف الأسماء ، اعتمدا على تحليل للأعمال الكتابية قام به الدكتور محمد عبد السلام احمد بمعاونة ديوان الموظفين عام ١٩٥٢ . وقنن الاختبار الأول على عينة حجمها ٨٢٢ مفحوصا من الجنسين من الحاصلين على شهادة الثانوية العامة او دبلوم التجارة الثانوية تتراوح أعمارهم بين ١٨ ، ٢٨ سنة ، وقنن الاختبار الثانى على عينة عددها ٨٨٠ فردا من الجنسين تتراوح أعمارهم أيضا بين ١٨ ، ٢٨ سنة من حملة الثانوية العامة المتقدمين لشغل الوظائف الكتابية عن طريق اعلانات ديوان الموظفين فى ذلك الوقت . وأعدت للاختبارين معايير مئينية وثانية وحسب لهما معامل الثبات بطريقة التجزئة النصفية فبلغ ٠.٩٧ (ن = ١٠٠) والواقع أن استخدام طريقة التجزئة النصفية فى حساب اختبارات القدرة الكتابية - التى تعتمد اعتمادا كليا على السرعة - يثير كثيرا من المشكلات أشرنا إليها فى الفصول السابقة إلا اذا كان المؤلف لجا الى تجزئة زمن الاختبار ، وهذا ما لم توضحه كراسة تعليمات هذه الاختبارات (١٢ : ٢١) . أما عن صدق الاختبارين فقد حسب مؤلفهما معاملات الارتباط بينهما وبين اختبارات المهن الكتابية والذكاء الاجتماعى (ن = ٤٧ من طالبات معهد السكرتارية بالعباسية) ف لوحظ أن الاختبارين يشتركان مع الاختبارات التى تقيس الجوانب الآتية : السرعة ، والدقة ، سهولة استخدام اللغة ، الألفة بالأعداد ، الذاكرة المباشرة ، المهارة اليدوية .

(٢) اختبارات المهن الكتابية : من اعداد محمد عماد الدين اسماعيل وسيد عبد الحميد مرسى (٢٦) وتقيس مكونات ثلاث هى : القدرة العددية ، السرعة والدقة ، الاستدلال اللغوى . وقد حسب ثبات الاختبارات الثلاث (ن = ١٠٠ من طالبات معهد السكرتارية) بطريقة التجزئة النصفية قبلت ٨٦ ، ٨٤ ، ٨٥ على التوالى . ومرة أخرى نثير مسألة استخدام الطريقة الملائمة لحساب ثبات الاختبارات التى تعتمد على السرعة . واستخدم فى حساب صدق الاختبارات تقديرات مادة المحاسبة والترجمة وتقديرات المدرسين والمشرفين وتراوحت معاملات الارتباط بين ٤٢ر ، ٤٥ر . وحسبت المعايير المئينية للاختبارات .

اختبارات القدرات الموسيقية :

يوجد في الوقت الحاضر عدد من الاختبارات التي تصلح لقياس القدرات الموسيقية ، قد يكون اقدمها ظهورا او اكثرها اهمية اختبارات سيشور التي تسمى : Seashore Measures of Musical Talent

وقد ظهرت في صورتها الأصلية عام ١٩١٩ ، واستمرت شائعة الاستخدام في الولايات المتحدة دون تعديل حتى عام ١٩٣٩ حيث عدلت في صورتها التي تستخدم حتى الآن . وقد كانت هذه الاختبارات نتيجة للبحوث الواسعة النطاق في ميدان سيكولوجية الموسيقى والتي أجريت بجامعة ايوا تحت اشراف كارل سيشور . ويتألف هذه البطارية في صورتها الحديثة من ٦ اختبارات هي : تمييز الأصوات pitch ، شدة الصوت loudness ، تذكر الايقاعات rhythm ، الزمن time ، نوعية الصوت timbre ، تذكر الألحان tonal memory .

ويتألف اختبار تمييز الأصوات من ٥٠ زوجا من النغمات ، وعلى المفحوص أن يحدد بالنسبة لكل زوج ما اذا كانت النغمة الثانية أحد ام اغلظ من النغمة الأولى وفي اختبار شدة الصوت يطلب منه أن يحدد بالنسبة لكل زوج من ٥٠ زوجا من النغمات ما اذا كانت النغمة الثانية اشد ام اضعف من الأولى . ويطلب منه في اختبار تذكر الايقاعات أن يحدد ما اذا كان النغمان الايقاعيان اللذان يتألف منهما كل سؤال (من ٣٠ سؤالا) متشابهين او مختلفين : وفي اختبار الزمن يحدد المفحوص ما اذا كانت النغمة الثانية اطول ام اقصر من النغمة الأولى (٥٠ سؤالا) . وفي اختبار نوعية الصوت يقوم المفحوص بالحكم على النغمة الثانية بانها متشابهة او مختلفة عن النغمة الأولى من حيث نوعية الصوت (البناء الهارموني) ، ويتكون هذا الاختبار من ٥٠ سؤالا . اما الاختبار الأخير ، وهو تذكر الألحان ، فيتكون من ٣٠ زوجا من النماذج اللحنية ، وفي كل زوج توجد نغمة واحدة مختلفة في النموذج الثاني ، وعلى المفحوص أن يحدد أي النغمات اختلفت ويحدد ترتيبها الرقمي .

وتصلح هذه الاختبارات للتطبيق على المفحوصين ابتداء من الصف

الرابع الابتدائي حتى أعلى المستويات التي تتطلب أعدادا موسيقيا رفيعا للمرافقين والراشدين .

وقد ظهرت هذه الاختبارات في أصلها الأجنبي مسجلة على أسطوانات، وقد أعدتها باللغة العربية أمال أحمد مختار صادق (٥) مسجلة على شرائط ، وأعدت تعليمات تطبيقها . كما صممت كراسة اجابة مختلفة عن الطبعة الأجنبية . وحسبت معاملات الثبات بطريقة التجزئة النصفية وتصحيح سبيرمان براون على عينة من ١٦٢ طالبا وطالبة من حملة الثانوية العامة المتقدمين للالتحاق بالمعهد العالي للتربية الموسيقية وتراوحت معاملات الثبات بين ٥٨ر (اختبار الزمن) ، ٨٨ر (اختبار تذكر الألحان) . أما عن صدق هذه الاختبارات فقد حسبت الباحثة معاملات الارتباط بين الاختبارات الفرعية التي تتألف منها البطارية لعينات من طلاب دور المعلمين والمعلمات (ن = ١٧١) وطلاب المعاهد الموسيقية العليا (ن = ٦٠) ، وتلاميذ المدارس الإعدادية العامة (ن = ١١١) ، وتلاميذ المعهد القومي للموسيقى (ن = ٧٤) . كما حسبت معاملات الارتباط بين اختبارات سيشور واختبارات القدرات الموسيقية للأطفال (بنثلي) لعينة تلاميذ المرحلة الإعدادية العامة . وتلاميذ المعهد القومي للموسيقا . وكذلك معاملات ارتباطها باختبارات ونج للذكاء الموسيقي، واريجون للتذوق الموسيقي، وفارنم للتدوين الموسيقي وكاتل للذكاء العام . واختبار الذكاء المصور للدكتور أحمد زكي صالح ، وتوافر في كراسة التعليمات بيانات كافية عن المعالم السيكمترية لهذه البطارية . أما عن معايير هذه البطارية فقد حسبت المئينيات الخاصة بكل اختبار فرعي ولكل عينة فرعية من عينة التقنين (٥) كما تؤكد الصدق التنبؤي لبطارية سيشور من خلال دراسة تتبعية قامت بها أمال صادق ، واکرام مطر (٧) استخدم فيها أسلوب تحليل الانحدار المتعدد وتأكد فيها ارتفاع القيمة التنبؤية لاختبارات سيشور في مجال القبول للتعليم الموسيقي في مصر بمقارنتها بالأساليب التقليدية التي تستخدم حتى وقتنا الحاضر في انتقاء الطلاب الجدد في كلية التربية الموسيقية جامعة حلوان .

وقد أعدت نفس الباحثة (٤) صورة مختصرة لبطارية سيشور ووصفت معالمها السيكمترية بوضوح .

ويوجد اختبار آخر للقدرة الموسيقية ظهر عام ١٩٦٢ أعده في إنجلترا

نولد بنتلى Bentley خصيصا لاطفال المرحلة الابتدائية ، (من ٧ - ١٦ سنة) و يقيس ٤ قدرات هي : تمييز الأصوات ، تذكر النغمات ، تحليل التالقات ، تذكر الايقاعات . ويمكن تطبيق هذه الاختبارات على مستويات عمرية أعلى مع مراعاة معايير هذه الأعمار عند تفسير الدرجات .

ويتكون اختبار تمييز الأصوات من ٢٠ زوجا من الأصوات ، النقية ، يستمع اليها المفحوص ثم يطلب منه أن يحدد ما اذا كان الصوت الثانى من كل زوج متشابها مع الصوت الأول أو مختلفا عنه . وفى حالة الاختلاف على المفحوص أن يحدد ما اذا كان أعلى (أى أحد) أو أسفل (أى أغلظ) من الصوت الأول . أما اختبار تذكر النغمات فيتكون من ١٠ أسئلة يتطلب كل منها مقارنة ثنائية بين جملتين لحنيتين قصيرتين ، ويتكون كل منها من ٥ نغمات ، ويطلب من المفحوص أن يحدد ما اذا كانت الجملة الثانية متشابهة مع الأولى أو مختلفة عنها وإذا حدث اختلاف عليه أن يحدد ترتيب النغمة المختلفة . وفى اختبار تحليل التالقات يطلب من المفحوص تحديد عدد النغمات التى يتكون منها كل تآلف (يتكون الاختبار من ٢٠ تالفا) . أما اختبار تذكر الايقاعات فيتكون من ١٠ مفردات يتطلب كل منها مقارنة ثنائية بين جملتين ايقاعيتين ، وتتكون كل واحدة من ٤ وحدات ايقاعية أو ما يعادلها من النماذج ايقاعية المختلفة ، ويطلب من المفحوص تحديد ما اذا كانت الجملة الثانية فى كل سؤال متشابهة مع الجملة الأولى أو مختلفة عنها . وفى حالة الاختلاف يكون عليه تحديد رقم الوحدة المختلفة .

وقد أعدت هذه الاختبارات باللغة العربية أيضا أمال أحمد مختار صادق عام ١٩٧٢ (٦) . وظهرت اختبارات القدرات الموسيقية للأطفال مسجلة على شرائط (بدلا من الاسطوانات فى الأصل الأجنبى) وأعدت لها كراسة اجابة خاصة : كما فصلت تعليمات تطبيقها . وقد استخدم المؤلف الأصل فى حساب ثبات هذه الاختبارات طريقة اعادة الاختبار ، وبلغ معامل الارتباط ٨٤ر . واستخدمت الباحثة المصرية هذه الاختبارات فى دراستين عامليتين للقدرة الموسيقية (٦) احدهما على عينة من تلاميذ المرحلة الاعدادية (ن = ١١١) ، والاخرى على عينة تناظرها فى العمر الزمنى من تلاميذ المعهد القومى للموسيقى (الكونسيرفاتورى) عددها ٧٤ تلميذا . وتراوحت اشتراكيات الاختبارات الفرعية بين ٤٢ر (لاختبار تذكر

الايقاعات لعينة تلاميذ المرحلة الاعدادية العامة) ، ٦٨ ر (لاختبار تمييز الأصوات لنفس العينة) . أما اشتراكيات الدرجة الكلية للبطارية فبلغت ٨٨ ر فى الحالتين .

ولحساب صدق الاختبارات حسبت معاملات الارتباط بين الاختبارات الفرعية ، ومعاملات ارتباطها باختبارات سيشور التى اشرنا اليها ، وباختبارات كاتل للذكاء العام . وكذلك حسبت المعايير الثنائية لعينتين من التلاميذ : احدهما مجموعة ٩ - ١٢ سنة (ن = ١٦٢) ، والاخرى مجموعة ١٢ - ١٥ سنة (ن = ٣٦٢) .

ويوجد اختبار آخر ظهر فى انجلترا عام ١٩٤١ ثم ظهرت طبعته الجديدة المعدلة عام ١٩٥٨ ثم عام ١٩٦٢ وهو اختبار ونج للذكاء الموسيقى .

Wing Standardized Tests of Musical Intelligence

ويصلح للتطبيق ابتداء من سن ٨ سنوات ، ويختلف عن اختبار سيشور فى انه لا يؤكد الجانب « الذرى » التحليلى كما يتمثل فى عناصر الاحساس البسيط السائد فى اختبارات سيشور ، ولذلك فانه يعتمد على محتسوى موسيقى واقعى حيث تستخدم موسيقى البيانو فى كل اختبار من الاختبارات السبع التى تتألف منها بطارية ونج وهى : تحليل التألفات ، تمييز الأصوات ، ذاكرة الأصوات ، التوافق الموسيقى (الهارمونى) ، الشدة ، الايقاع ، الجمل الموسيقية . وتتطلب الاختبارات الثلاث الأولى تمييزا حسيا على مستوى اكثر تعقدا من اختبارات سيشور ، أما فى الاختبارات الأربع الأخرى فإنّ المفحوص يقارن النوعية الجمالية . مقطوعتين موسيقيتين ، ومعنى ذلك ان بطارية ونج تؤكد جانب التدقيق الموسيقى . ويقترح ونج استخدام الاختبارات الثلاث الأولى كبطارية مختصرة مع الاطفال الصغار او كأداة تصفية مبدئية .

وقد نشرت آمال أحمد مختار صادق (٢) فى عام ١٩٧٩ تقريرها عن تقنين هذا الاختبار على البيئة المصرية . وقد بلغت عينة التقنين ٢٤١ مفحوصا من الذكور والإناث اشتقوا من مستويات تعليمية موسيقية مختلفة من دور المعلمين والعلماء والمعهد القومى للموسيقى (الكونسيرفتوار) .

مؤكلية التربية الموسيقية بجامعة حلوان - وقد سارت الباحثة فى تقدير صدق المحتوى من خلال استطلاع رأى المهتمين بالموسيقى بدرجة « الموسيقية » فى هذه الاختبارات . أما بالنسبة لصدق التكوين الفرض بحساب معاملات الارتباط بين الاختبارات الفرعية للبطارية وتحليل الفروق من مستويات التدريب الموسيقى وتمايز المستويات التعليمية الموسيقية ، كما استخدم الصدق المرتبط بالمحكات باستخدام بطارية سيشور التى سبق تفنيها من ناحية والتحصيل الموسيقى من ناحية أخرى ، وكانت هذه النتائج جميعا مدعمة لصدق البطارية .

أما الثبات فقد حسب معاديتين : التجزئة النصفية ، كيبور - ريتشارد سمون . وحسبت المعايير المئينية للاختبارات الفرعية وللمجموع الاختبارات الإدراكية ومجموع الاختبارات التذوقية والدرجة الكلية فى البطارية . كما حسبت أيضا نسب الذكاء الانحرافية التى أطلقت عليها الباحثة تسمية « نسب الذكاء الموسيقى » ومصطلح « الذكاء الموسيقى » هو المصطلح الأثير عند ونج مؤلف هذه الاختبارات .

وفى عام ١٩٧٩ نشرت أعمال صادق أيضا دراسة هامة تتضمن عرضا مفصلا لبطارية جديدة لقياس القدرات الموسيقية قامت بإنشائها (٣) فى محاولة منها للارتباط بين مجال قياس القدرة الموسيقية والتقدم فى مجال دراسة القدرات العقلية بصفة عامة وما أحدثته نتائج البحوث فيها من تغير لم يظهر له أثر فى بطاريات سيشور ونج وبنقلى التى سبق الإشارة إليها . وتشمل الاختبارات الجديدة التى قامت الباحثة بإنشائها ما يلى :

(١) اختبار الإغلاق الموسيقى : musical closure ويتألف من قسمين أولهما من ٢٤ مفردة فى كل منها يستمع المتحوص الى جملة موسيقية قصيرة ، وبعدها بفترة يستمع الى جملة موسيقية أكبر ، ويطلب منه تحديد ما اذا كانت الجملة الأولى متضمنة فى الجملة الثانية أم لا . وثانيهما ويتألف من ١٢ مفردة ، كل منها عبارة عن جملة موسيقية اختيرت من مؤلفات للبيانو ، ويعرض على المتحوص (موتيف) وعليه أن يقرر ما اذا كان موجودا فى اللحن أم لا .

(٢) اختبار تعيين الألحان : melody identification ويتألف من صورتين أحدهما للمراهقين والراشدين والثانية للأطفال . وتتكون صورة الكبار من قسمين أولهما يتألف من ٤٨ مفردة من الألحان فى السلام الكبيرة والصغيرة الغربية ، وثانيهما من ١٢ مفردة من الألحان التى تتضمن المسافة العربية (أى تمثل الموسيقى العربية) . أما اختبار الأطفال فيتألف من ٢٠ مفردة كنها من نوع الموسيقى الغربية . وفى جميع الحالات يقوم المفحوص بالمقارنة بين المعزوف والمدون على هيئة نوتة موسيقية .

(٢) اختبار التعيين الإيقاعى : rhythmic identification وقد أعد من صورتين أيضا . وتتألف صورة المراهقين والراشدين من ٦٠ مفردة ، عشرة منها فى كل ميزان من الموازين الموسيقية ويطلب من المفحوص أن يقارن نمطا إيقاعيا منقرا بنمط إيقاعى مكتوب على هيئة تدوين موسيقى ، أما صورة الأطفال فتتألف من ٢٠ مفردة كلها فى الميزان الموسيقى البسيط .

(٤) اختبار التعرف على المسافات : interval recognition ويتألف من صورة واحدة تصلح للأطفال لسهولة الظاهرة بالنسبة للمراهقين والراشدين ، وينقسم الى قسمين : أولهما يتكون من ٢٠ زوجا من المسافات تعزف عزفا متتابعيا على البيانو باستخدام طبقات صوتية مختلفة ، ويطلب من المفحوص أن يقرر ما إذا كانت المسافتان متشابهتين أو مختلفتين . وفى القسم الثانى : مفردات أكثر صعوبة تعزف عزفا متتابعيا على الكمان .

(٥) اختبار تقدير المسافات : interval estimation ويتألف من صورة واحدة للمراهقين والراشدين لصعوبته الظاهرة بالنسبة للأطفال ويتألف من ٤٠ مفردة مصنفة الى ثلاثة أقسام تتناول على التوالى المسافات الصاعدة فى الموسيقى الغربية ، والمسافات الهابطة منها ، ومسافات الموسيقى العربية . وفى جميع الأقسام تعزف فى المفردة الواحدة مسافة مرجعية أصلية مرة واحدة على البيانو أو الكمان وبعد فترة توقف طوله ٥ ثوان تعزف ثلاث مسافات على التوالى ويطلب من المفحوص أن يحكم على أى هذه المسافات الثلاث هى الأقرب الى المسافة الأصلية .

وقد قننت البطارية على عينة مصرية تتألف من تلاميذ المرحلة الإعدادية العامة (ن = ١١١) ، وتلاميذ المرحلتين الابتدائية والإعدادية بالمعهد القومى

للموسيقى (ن = ٧٤) ، وطلاب الكليات والمعاهد الموسيقية (ن = ٦٠) .
وحسب الثبات بطريقة كيودر - ريتشارد سون ، واستخدمت في حساب
الصدق محكات تتمثل في الاختبارات التي تقيس القدرة الموسيقية والتي سبق
تقنينها على البيئة المصرية (ميشور ، بنتلى ، ونج) كما استخدم محك
التحصيل الموسيقى ، وكانت المؤشرات الاحصائية لكل من الثبات والصدق
في صالح البطارية . كما استخدمت الباحثة طريقة أخرى في حساب
الصدق تتمثل في المقارنة بين العينات ذات المستويات المختلفة من الانتقائية
من كل من الأطفال والمراهقين والراشدين وكانت النتائج في معظم الحالات
دالة لصالح المجموعات الأعلى انتقائية (تلاميذ الكونسيرفتوار من الأطفال
وطلاب المعاهد والكليات الموسيقية من المراهقين والراشدين) .

اختبارات القدرة في الفنون البصرية :

توجد في الوقت الحاضر طائفة ن الاختبارات التي تقيس بعض جوانب
القدرة في ميدان الفنون البصرية ، ومنها اختبار Graves Design
Judgment Test الذي يهدف الى قياس قدرة المفحوص على ادراك المبادئ
الأساسية للعمل الجمالي والاستجابة لها وهي مبادئ الوحدة unity
والسيادة dominance والتنوع variety والتوازن balance
والاستمرار continuity والتناسق symmetry والتناسب
proportion والایقاع rhythm . وقد تحقق ذلك بعرض
٩٠ ثنائية وثلاثية من الرسوم المجردة ، في كل منها واحد من الرسوم
يتفق مع المبادئ الثماني هذه بينما يتمثل أحد هذه المبادئ او بعضها في
الرسم الآخر أو الرسوم الأخرى . ويطلب من المفحوص ان يفاضل بين رسوم
الثنائية أو الثلاثية ويختار أحدها على اساس التفضيل الفني . ولم يتضمن
هذا الاختبار نماذج من أعمال الفن ، لأن الحكم على الصور واللوحات يتأثر
بالخبرة السابقة وسمات الشخصية والانفعالات ومفهوم المفحوص عن الفن .
ولذلك فان الاختبار يؤكد محض الاختبار الجمالي ، او ما يسميه احد مؤلفي
هذا الكاب « الحساسية الجمالية » (٢٠) . وبالإضافة الى ذلك فان الاختبار
يقيس العمليات الإدراكية المعقدة وليس العناصر الحسية البسيطة التي تتدخل
في الحكم الجمالي ، وفي هذا يتشابه مع اختبارات ونج في الموسيقى ، ومع
المبادئ المتضمنة في قياس العامل العام في الذكاء .

والاختبار الثانى فى هذا المجال هو اختبار ماير للحكم الفنى .
Meier Art Judgment Test

ويهدف الى قياس الحكم الجمالى بطريقة « كلية » ، وبخلاف من اختبار جريفز فى انه يتألف من ١٠٠ زوج من اللوحات الفنية مرسومة بالأبيض والأسود ، ويمثل أحد عنصرى الثنائية عملاً فنياً ممتازاً ، والعنصر الآخر أدخل عليه بعض التفسير والتعديل فى أحد الجوانب الفنية الهامة تجعله أقل جودة من العنصر الأول . ويحدد المفحوص أى الجوانب قد تغير (الشكل أو الزوايا مثلاً) إلا انه لا يعلم أى اللوحتين هى اللوحة الأصلية ، ويطلب من المفحوص أن يحدد تفضيله لأى منهما .

ويرى ماير أن هذا الاختبار يقيس الحكم الجمالى ، إلا انه فى رأى أحد مؤلفى هذا الكتاب يقيس « الحساسية الجمالية » أيضاً (٢٠ : ٥) ، والفرق بينه وبين اختبار جريفز فرق فى المحتوى وليس فى العمليات الأساسية المتضمنة .

وإذا كان اختبار جريفز وماير ينتميان الى ميدان التذوق الفنى فإن اختبارى كنوبر Knauber وهورن Horn ينتميان الى ميدان الابتكار الفنى . ويسمى أولهما Knauber Art Ability Test ويصلح للمراهقين والراشدين ، ويتطلب من المفحوص ما يأتى :

١ - أن يرسم من الذاكرة أشكالاً فى حدود المكان المسموح به .

٢ - أن يرسم شخصيات نمطية (بابا نويل مثلاً) .

٣ - ترتيب مؤلف معين فى حدود المكان المسموح به .

٤ - ابتكار وتكملة رسوم معينة من عناصر معطاة .

٥ - تحديد الأخطاء فى أعمال مرسومة (مثل المنظور الخاطئ) .

٦ - إنتاج أعمال فنية تدل على الخيال والحدق والتمثيل الرمزى .

١ أما اختبار هورن ويسمى Horn Art Aptitude Inventory فقد أعد خصيصا للاستخدام فى أغراض قبول طلاب كليات الفنون الجميلة ، ويقس على وجه الخصوص الابتكارية فى الفن مستخدما فى كل سؤال وصفا بسيطا مرنا أو مرشدا لكل رسم . ويتألف الاختبار من قسمين : الأول يتطلب من المفحوص أن يرسم ٢٠ شيئا مألوفا (منزل - كتاب ، الخ) فى وقت قصير ، كما يتطلب منه انتاج رسوم تجريدية بسيطة مبتدئا بمجموعة من المثلثات والمربعات وغيرها . أما القسم الثانى فيتطلب مستوى أعلى من الابتكارية وفيه يعطى للمفحوص ٢٠ بطاقة مستطيلة على كل منها خطوط تفيد كبداية لصورة ينتجها المفحوص . ويحكم على الرسوم على أساس الخيال الابتكارى من ناحية والخصائص الفنية من ناحية أخرى .

ومن الاختبارات المصرية التى تقيس القدرات فى الفنون البصرية اختبارات تكميل الأشكال وتكميل الصور والتقدير الجمالى التى أعدها الدكتور محمد عماد الدين اسماعيل (٢٥) لقياس عاملى الطلاقة والتقدير الجمالى . ويقس اختبار تكميل الأشكال عامل الطلاقة وفيه يطلب من المفحوص أن يضيف الى خطوط مرسومة بضعة خطوط أخرى بحيث تجعل منه صورة لشيء مألوف . ويقس اختبار تكميل الصور عامل الطلاقة أيضا وفيه يطلب من المفحوص أن يكتب فى مكان معين من صورة مرسومة أكبر عدد من الأشياء التى يمكن رسمها . أما اختبار التقدير الجمالى فهو تقنين مصرى لاختبار ماير الذى اشرنا اليه . وقد حسب صدق الاختبارات بإيجاز معاملات الارتباط بين كل منها وبين المحك (وهو تقدير المدرسين للقدرة الفنية) فكانت النتائج كالتالى : ٤٤ر لاختبار تكميل الأشكال ، ٢٤ر لاختبار تكميل الصور . ٤٤ر لاختبار التقدير الجمالى (ماير) . أما عن الثبات فقد حسب بمعادلة كيودر - ريتشاردسون وكان معامل ثبات اختبار تكميل الأشكال ٥٥ر ، واختبار تكميل الصور ٨٢ر ، واختبار (ماير) للتقدير الجمالى ٥٤ر .

وحسبت معايير الدرجات التائية للاختبارات الثلاثة ، وكانت عينة التقنين فى كل الأحوال ٦٥ طالبا من المعهد العالى للتربية الفنية .

اختيارات التفكير الابتكاري :

من أهم خصائص التقويم النفس المعاصر زيادة الاهتمام بقياس الابتكارية مسايرة للنشاط المتزايد في مختلف ميادين علم النفس لدراسة طبيعة النشاط الابتكاري والعوامل المؤثرة فيه . ولا يتسع المقام للافاضة في هذه المسائل، ويمكن الرجوع الى بعض المصادر المتاحة في اللغة العربية (١٩) ونستطيع القول بايجاز أن الاهتمام بالتفكير الابتكاري يعبر عن حاجة المجتمعات المعاصرة لزيادة وتنمية ثروتها البشرية من العلماء والمهندسين والقادة الاداريين والفنانين . ويرى كثير من العلماء المهتمين بهذا الميدان أن الابتكار عملية أساسية في هذه الأحوال جميعا ، فهو كما يقول تايلر « يختلف في العمق والمجال وليس في النوع » ومن رايه عدم التمييز بين الابتكار العلمي والابتكار الفني لأن الابتكار « يتضمن نظرة للمشكلات أكثر أساسية من التدريب المهني » .

ويعود الفضل الى العالم الأمريكي المعاصر جيلفورد في زيادة حركة قياس التفكير الابتكاري . فقد أثرت نظريته في تنظيم العقل في معظم بحوث الابتكار ، وفيها يقترح تصنيف القدرات العقلية الى فئتين كبيرتين هي قدرات التفكير ، وقدرات الذاكرة . وتنقسم عوامل التفكير الى ثلاثة أنواع هي : قدرات التفكير المعرفي (بمعنى الاكتشاف) ، وقدرات التفكير الانتاجي وقدرات التفكير التقويمي . ثم تنقسم قدرات التفكير الانتاجي الى فئتين فرعيتين هما : قدرات التفكير الانتاجي التقاربي ، وقدرات التفكير الانتاجي التباعدى . والانتاج التقاربي (أو الاختزالي) يتضمن تضيق الاحتمالات عند انتاج اجابة واحدة محتلة للمشكلة ، مثل : طويل الى قصير مثل مرتفع الى ٠٠٠ ؟ وبالطبع فان أغلب اختبارات الذكاء العام تقيس هذا النمط من التفكير ، بل ان معظم الاختبارات العقلية التي اشرنا اليها في الفصول السابقة من هذا الكتاب تنتمي الى فئة الانتاج التقاربي . أما التفكير التباعدى فيتطلب انتاج اكبر عدد ممكن من الاجابات . مثل : ما هي الاستخدامات التي يمكن أن تفكر فيها لقلب الطوب ؟ » .

ويرى جيلفورد ان الانتاج التباعدى عامل هام في التفكير الابتكاري .
واستخدم في وضع اختبارات عديدة لقياس هذه القدرة في مستويات مختلفة

وفى محتويات متباينة • ومن رايه ان بعض عوامل الذاكرة والتفكير المعرفى والتفكير التقويمى والتفكير الانتاجى التقارىى ترتبط بالعمل الابتكارى • ويعتقد على وجه الخصوص ان الموهبة الابتكارية يمكن قياسها بالاختبارات التى تقيس قدرات : الأصالة والطلاقة والمرونة من مختلف الأنواع ، وقدرات إعادة التحديد والتضمين والتفكير التقويمى أيضا •

وقد اثرت نظرية جيلفورد فى ظهور بطاريتين رئيسيتين لقياس التفكير الابتكارى احدهما أعدهما جيلفورد وزملاؤه بجامعة كاليفورنيا • والأخرى أعدهما تورنس بجامعة مينيسوتا •

اختبارات جيلفورد للتفكير الابتكارى :

تتألف بطارية اختبارات التفكير الابتكارى لجيلفورد على النحو الآتى :

١ - اختبارات جيلفورد - كريستنسن للطلاقة وتتألف من ٤ اختبارات

(١) طلاقة الكلمات (الانتاج التباعدى لوحداث الرموز) : يطلب فيه من المفحوص ان يكتب كلمات ذات خصائص معينة ، كأن تتضمن حروفا معينة (ب مثلا) •

(ب) طلاقة الأفكار (الانتاج التباعدى لوحداث المعانى) : يطلب فيه من المفحوص ان يذكر أسماء الأشياء التى تنتمى الى فئة معينة (مثل السوائل التى تحترق) •

(ج) طلاقة التداعى (الانتاج التباعدى للعلاقات بين المعانى) • يطلب فيه من المفحوص اعطاء كلمات تتشابه فى المعنى مع كلمة أخرى •

(د) الطلاقة التعبيرية (الانتاج التباعدى لمنظومات الرموز) : يطلب فيه من المفحوص تكملة جمل ، كل منها يتألف من ٤ كلمات ، بحيث تبدأ كل كلمة بحرف معين •

٢ - اختبار الاستخدامات البديلة (الانتاج التباعدى لفئات المعانى) :

يطلب فيه من المفحوص أن يذكر الاستخدامات المحتملة لشيء معين غير استخدامه المألوف (قالب الطوب مثلا يستخدم في البناء استخداما مألوفاً) .

٣ - اختبار المترقيات (الانتاج التباعدى لوحداث المعانى ، والانتاج التباعدى لتحويلات المعانى) . وفيه يطلب من المفحوص أن يذكر النتائج المختلفة المترتبة عن حادث افتراضى (مثل : ماذا يحدث لو أن الناس توقفوا عن النوم ؟) . ويعطى للمفحوص فى هذا الاختبار درجتان : احدهما للاستجابات الواضحة (وتقيس طلاقة الأفكار أو الانتاج التباعدى لوحداث المعانى) ، والاخرى للاستجابات البعيدة (وتقيس الأصالة ، والانتاج التباعدى لتحويلات المعانى) .

٤ - الأعمال المحتملة (الانتاج التباعدى لتضمينات المعانى) . ويطلب فيه من المفحوص أن يذكر الأعمال المحتملة التى يمكن أن يرمز لها بشعار ما (مصباح كهربائى مثلا) .

٥ - عمل الأشياء (الانتاج التباعدى لمنظومات الاشكال) : ويطلب فيه من المفحوص أن يرسم أشياء محددة باستخدام مجموعة معطاة من الأشكال (كالدوائر أو المثلثات) . وعلى المفحوص ألا يستخدم إلا الاشكال المعطاة فقط .

٦ - الاسكتشات (الانتاج التباعدى لوحداث الأشكال) : ويشبه اختبار عمل الأشياء فيما عدا أن كل صفحة فى الاختبار تشتمل على نفس الشكل (دوائر مثلا فقط) ويطلب من المفحوص أن يرسم أكبر عدد من الاسكتشات بإضافة تفاصيل الى كل شكل من الاشكال المعطاة .

٧ - مشكلات عيدان الثقاب (الانتاج التباعدى لتحويلات الأشكال) : وفيه يطلب من المفحوص أن يستبعد عددا معينا من عيدان الثقاب (المرسومة على الورق فى صورة خطوط مستقيمة) بحيث ينتج عن ذلك عددا من المربعات أو المثلثات .

٨ - اختبار الزخرفة (الانتاج التباعدى لتضمينات الاشكال) : وفيه (التقويم النفسى)

يطلب من المفحوص أن يزخرف صورا لأشياء مألوفة بأكبر عدد من الرسوم المختلفة .

ويلاحظ أن الاختبارات السبع الأولى من هذه القائمة تتطلب استجابات لفظية ، بينما تستخدم الاختبارات الأربع الأخيرة محتوى من الأشكال أو الصور . وهذه الاختبارات الأحد عشر تمثل عينة صغيرة من مجموعات الاختبارات التي أنتجتها بحوث جيلفورد وزملائه وتلاميذه بجامعة جنوب كاليفورنيا لأكثر من عشرين عاما . وقد أعدت معايير مبدئية لهذه الاختبارات في صورة مئينيات ومعايير جيمية C-Score (وهو معيار يلخص المستويات التائية الى ١١ مستوى وهو في ذلك يشبه المعيار التساعي ويختلف عنه في عدد المستويات) . وفي معظم الأحوال اشتقت هذه المعايير لعينات من الراشدين أو طلاب من سن ١٥ + أو كليهما . وتصلح هذه الاختبارات بوجه عام لمستوى المرحلة الثانوية وما بعدها ، رغم أن جيلفورد يذكر أن بعض اختبارات تصلح لمستوى المرحلة الابتدائية (ابتداء من الصف الأول) .

والمشكلة الجوهرية في اختبارات التفكير الابتكاري هي مشكلة التصحيح فهي تتطلب من المصحح تدريباً كافياً على اتباع تعليمات التصحيح ودراسة الأمثلة التي تتضمنها كراسات التعليمات . ورغم ذلك فإن عملية تصحيح هذه الاختبارات تستغرق وقتاً كبيراً .

وحسب جيلفورد ثبات هذه الاختبارات بطريقة التجزئة التصفية . وتتراوح معاملات الثبات بين ٠.٦٠ ، ٠.٨٠ . أما الصدق فقد اعتمد اعتماداً كبيراً على صدق التكوين الفرضي كما يتمثل في اتفاق الاختبارات مع نموذج العقل الذي يقترحه . ولا زالت هذه الاختبارات في حاجة إلى أنواع الصدق الأخرى وخاصة الصدق المرتبط بالمحكات .

وقد قام عبد السلام عبد الغفار (١١) بأعداد بعض اختبارات جيلفورد باللغة العربية ، فأعد اختبارين لقياس الطلاقة اللفظية (طلاقة الكلمات) هما اختبار الطلاقة اللفظية ١ ، ويتطلب من المفحوص إنتاج كلمات تبدأ

بحرف معين ، واختيار للطلاقة اللفظية ٢ ويتطلب من المفحوص أن ينتج كلمات تنتهي بحرف معين . كما أعيد ٣ اختبارات أخرى لقياس بعض عوامل التفكير الابتكاري الأخرى وهي : اختيار الطلاقة الفكرية (لقياس عامل طلاقة الأفكار) ، واختيار الاستعمالات (لقياس عامل المرونة التلقائية ، واختبار المترتبات (لقياس عامل الأصالة) . وقد حسب ثبات هذه الاختبارات من أداء عينة عددها ١٢٠ تلميذا بالمدارس الثانوية العامة بطريقة إعادة الاختبار لاختبار الطلاقة اللفظية فبلغ معاملات الثبات ٠٦٦ ، ٠٦٢ ، ٠٦٦ ، ٠٦٩ ، ٠٧٥ ، ٠٨٠ ، على التوالي . وفي تقدير الصدق حسب الباحث معاملات ارتباط الاختبارات بالتحصيل المدرسي ، وتراوحت هذه المعاملات بين ٠٢٧ (الاستعمالات) و ٠٥٨ (المترتبات) ، ولم تتضمن كراسة التعليمات بيانات عن معايير الاختبارات .

كما قام سيد خير الله (١٠) بأعداد بعض اختبارات جيلفورد أيضا
 وصدر له اختبار القدرة على التفكير الابتكاري يتكون من خمسة أقسام
 تقيس : الاستمالات والمرتبات ولعبة الكلمات ، ولا تتوافر لدينا أى معلومات
 سيكومترية عن هذا الاختبار .

اختبارات تورنس للتفكير الابتكاري :

أشرنا الى أن اختبارات جلفورد هي محصلة بحوثه العملية في ميدان التفكير الانساني وطبيعة الذكاء ، أما اختبارات تورنس للتفكير الابتكاري التي نتناولها الآن فقد ظهرت لمواجهة الضرورات التربوية والعملية كجزء من برنامج طويل من البحوث تهتم بالخبرات التعليمية التي تساعد على تنمية الابتكار . وبعض اختبارات تورنس هي في الواقع تعديل للاساليب التي استخدمها جيلفورد . بالإضافة الى أن الدرجات المشتقة من البطاريات الكاملة تعتمد على العوامل التي توصلت اليها بحوث جيلفورد وهي الطلاقة والمرونة والأصالة واعطاء التفاصيل . وفي هذا الصدد يذكر تورنس أنه لم يسع الى وضع اختبارات « نقية عامليا » أو « بسيطة التكوين » ، وإنما الى التوصل الى اختبارات تتضمن مواقف مماثلة لما تتطلبه العملية الابتكارية

فى الظروف الطبيعية والمعقدة . ولذلك فان كل اختبار قد يصحح فى ضوء عاملين ، او ثلاثة عوامل ، او العوامل الأربع جميعا .

وتتألف اختبارات تورنس للتفكير الابتكارى من ١٠ اختبارات ، مصنفة الى بطاريتين : احدهما لفظية والاخرى مصورة . وتسمى البطارية الأولى : التفكير الابتكارى بالكلمات ، والثانية : التفكير الابتكارى بالصور . وحتى يمكن التخفيف من التهديد الذى تتضمنه كلمة « اختبار » استخدمت فى كل الأحوال كلمة « نشاط » كما تؤكد التعليمات على « اللعب » ، وتوصف هذه الاختبارات بأنها تصلح لمختلف المستويات ابتداء من مرحلة الحضانة حتى مستوى الدراسات العليا ، بشرط أن تطبق فرديا وشفويا على المستويات الأدنى من الصف الرابع الابتدائى . وتوجد صورتان متكافئتان لكل بطارية من البطارتين .

وفى « التفكير الابتكارى بالكلمات » الصورة ا والصورة ب تتألف الأنشطة الثلاث الأولى (اسأل وخمن) من صورة غير عادية يستجيب لها المفحوص بكتابه (١) جميع الأسئلة التى يحتاج الى توجيهها لكى يعرف ما يحدث فى الصورة (ب) الأسباب المحتملة لما حدث فى الصورة (ج) النتائج المترتبة على ما حدث فى الصورة . أما النشاط الرابع فيتعلق بطرق تحسين لعبة من لعب الأطفال بحيث تصبح مصدرا لمزيد من السرور والفرح لمن يلعب بها . ويتعلق النشاط الخامس بذكر الاستخدامات غير العادية لشيء مألوف : (لعب الكرتون) فى الصورة (ا) ، (وعلب الصفيح) فى الصورة (ب) . ويتعلق النشاط السادس بالأسئلة غير المعتادة التى يمكن توجيهها عن الشيء المتضمن فى النشاط الخامس . أما النشاط السابع فقد أعد على نسق اختبار الترتيبات عند جيلفورد ، ويتطلب من المفحوص ان يذكر كل ما يمكن ان يحدث اذا نشأ موقف غير ممكن الحدوث . وتعطى البطارية درجات كلية فى العوامل الثلاث : الطلاقة والمرونة والأصالة .

ويتألف اختبار « التفكير الابتكارى بالصور » : الصورة (١) والصورة (ب) ، من ثلاثة أنشطة :

(١) تكوين الصورة وفيه يعطى للمفحوص ورقة ملونة ذات شكل منحنى وعليه ان يضعها على صمحة بيضاء فى الموضع الذى يشاء ويستخدمها

كبداية لرسم صورة غير عادية (والورقة الملونة المستخدمة فى الصورة
(١) زرقاء اللون ، والمستخدم فى الصورة (ب) صفراء اللون فى الطبعة
العربية .

ويحصل المفحوص فى البطارية المصورة على ٤ درجات هى : الطلاقة
والمرونة والأصالة والتفاصيل .

وقد أعد هذه البطاريات باللغة العربية (بموافقة مؤلفها تورنس)
عبد الله محمود سليمان وفؤاد أبو حطب (٦) وظهرت عام ١٩٧١ على النحو
الآتى :

- ١ - التفكير الابتكارى بالكلمات : الصورة ١ .
- ٢ - التفكير الابتكارى بالكلمات : الصورة ب .
- ٣ - التفكير الابتكارى بالصور : الصورة ١ .
- ٤ - التفكير الابتكارى بالصور . الصورة ب .
- ٥ - كراسة تعليمات اختبارات تورنس للتفكير الابتكارى .
- ٦ - استمارة تصحيح (١) لاختبارات التفكير الابتكارى بالكلمات
(الصورة (١) أو الصورة (ب)) .
- ٧ - استمارة تصحيح (٢) لاختبارات التفكير الابتكارى بالصور
(الصورة (١) والصورة (ب)) .
- ٨ - استمارة تقدير المعلمين للابتكار .
- ٩ - استمارة تقدير التلاميذ للابتكار .

وفى عام ١٩٧٧ نشر معدا الاختبار تقريرهما الأول (١٦) عن تقنين هذه
البطارية ، ويشمل على وجه الخصوص تقنين الصورة ١ من اختبارات الأشكال
أو الصور (*) ، فطبق الاختبار على ٣١٢ تلميذا وتلميذة من مدرستين اعداديتين
بالقاهرة تتراوح اعمارهم بين ١٢ ، ١٦ سنة ، ولحساب الصدق استخدمت

(*) عاون الباحثين فى تصحيح الاختبارات التى طبقت فى البحث المشار اليه
يومئذ السادة (الدكتور فىما بعد) محمود السيد عمر ، على السيد سليمان ، سيد
محمود الطراب ، محمد متولى غنيمه ، والاستاذ محمد السيد عبد الحليم . وبعد
انقضاء اكثر من ١٥ عاما على مشاركتهم الكريمة فزجى اليهم الشكر والاعتراف
بالفضل .

استمارة تقدير المعلمين للابتكار حيث طلب من مدرسي الفرق التي أجريت عليها الاختبارات أن يختار كل منهم ٥ تلاميذ يقومون في الفئة العليا ، ٥ تلاميذ يقومون في الفئة الدنيا في كل من أبعاد الابتكار المقيسة وهي : الطلاقة ، المرونة ، الأصالة ، التفاصيل . واختير كل تلميذ يتفق عليه مدرسان على الأقل . وحسبت دلالة الفروق بين متوسطات مجموعات التلاميذ الأعلى والأدنى في مختلف أبعاد الاختبار باستخدام مقياس (ت) فوجدت الفروق في معظم الأحوال (باستثناء التفاصيل) دالة وفي الاتجاه الصحيح إلا أن مستوى الدلالة كان عاليا (٠.١) بالنسبة إلى بعد المرونة ، ومنخفضا نسبيا (٠.١) بالنسبة إلى بعدى الطلاقة والأصالة .

وقد استخدم في حساب الثبات الطرق الآتية .

١ - حساب معاملات الارتباط بين متغيرات الاختبار (الطلاقة ، المرونة ، الأصالة ، التفاصيل ، الدرجة الكلية) فوجدت المعاملات في كل الأحوال عالية وذات دلالة احصائية .

٢ - حساب معاملات الارتباط بين مصححين قاما بتصحيح ٣٨ اختبارا كل على حدة ، وكانت معاملات ثبات التصحيح ٠.٩٩ ، للطلاقة ، ٠.٩٩ ، للمرونة ، ٠.٩٩ ، للأصالة ، ٠.٩٧ ، للتفاصيل ، ٠.٩٩ ، للدرجة الكلية . ويرجع ارتفاع المعاملات على هذا النحو إلى التزام المصححين بقواعد مقننة في التصحيح ، كما لجا الباحثان في تدريب المصححين على طريقة تورنس ، وتأكدت بالفعل جدواها .

٣ - حساب الثبات بطريقة إعادة الاختبار (ن = ٣٨) وبلغت المعاملات ٠.٥٤ للطلاقة ، ٠.٣٦ للمرونة ، ٠.٥٣ ، للأصالة ، ٠.٤٠ ، للتفاصيل ، ٠.٥٠ ، للدرجة الكلية في الاختبار .

وقد أعد الباحثان معايير الاختبار في صورة درجات تائية ، كما أعدا النماذج التجريبية لتصحيح الصورة (١) من اختبارات الأشكال ، وهي نماذج مشتقة من استجابات عينة التقنيين المصرية ذاتها وليست منقولة عن كراسات التعليمات الأجنبية .

ولسوء الحظ لم تتوافر لمعدى الاختبار باللغة العربية فرصة أخرى للمعكوف على الاختبارات الثلاثة الأخرى فى البطارية لتقنينها ، كما لم تتح لهما الفرصة لاعادة تقنين اختبار الأشكال (الصورة ١) الذى مر على تقنيه الأول الآن أكثر من خمس عشرة سنة، الا ان مما يخفف من رقع الأمر ان الاختبارات الأربعة استخدمت فى بعض البحوث الامبيريقية وفى عدد من رسائل الماجستير والدكتوراه فى الجامعات المصرية وعدد من جامعات العالم العربى تؤكد نتائجها صلاحيتها للاستخدام .

اختبار التفكير الابتكارى للأطفال :

اعدت هذا الاختبار فى أصله الأجنبى سيلفيا ريم S. Rimm عام ١٩٧٦ بهدف قياس التفكير الابتكارى عند الأطفال ويطلق عليه اسم بطارية الاختبارات الجماعية للكشف عن الموهوبين Group Inventory for Finding Talent (GIFT)

ويتكون من جزئين رئيسيين : اولهما ويصلح للأطفال من سن ٦ - ٩ سنوات ، وثانيهما ويصلح للأطفال من سن ٩ - ١٢ سنة ويتضمن جزءا مصورا من اختبارات تورنس (الدوائر والمربعات) وجزءا لفظيا من اختبارات جيلفورد (الاستعمالات) ، بالإضافة الى جزء يتألف من عبارات تتضمن صفات الطفل المبتكر .

وقد اعد الاختبار باللغة العربية سيد خير الله ومحمود عبد الحليم منسى (١١) ، وقاما بتقنيه على عينة من ٥٠٠ طفل من الذكور والاناث من تلاميذ المدارس الابتدائية بمدينة الاسكندرية . وحسب الثبات بطريقة اعادة الاختبار وتراوحت معاملاته لمختلف صفوف المرحلة الابتدائية بين ٥١ر ، ٨٢ر . وفى حساب الصدق استخدم الباحثان تقديرات المعلمين للتفكير الابتكارى عند تلاميذهم كمحك ، كما استخدم ايضا اختبار القدرة على التفكير الابتكارى من اعداد سيد خير الله كمحك آخر .

اختبار التفكير الابتكارى عند الأطفال باستخدام الحركات والأفعال :

اعد هذا الاختبار فى أصله الاجنبى تورنس Torrance وجبس

Gibbs عام ١٩٧٧ ويعتمد في جوهره على استخدام الحركات والأفعال ليكون مناسباً للأطفال بين ٢ - ٧ سنوات . ويتناول أربعة أنشطة على وجه الخصوص هي : كم طريقة ؟ ، وتقدر متحرك مثل ؟ ، هل فيه طريقة أخرى ؟ الاستعمالات . وتقدر لهذه الأنشطة درجات للطلاقة والنخيل والأصالة .

وقد أعد الاختبار باللغة العربية محمد ثابت على الدين (٢٣) ، وأعد له معايير تصحيح ملائمة للعينات المصرية . وطبقه على مجموعات من الأطفال المقيدين بدور الحضانة وبعض المدارس الابتدائية بالمنصورة عام ١٩٨١ ، وحسب له الثبات بطريقة إعادة الاختبار . وكانت المعاملات مرضية ، وبذل الباحث محاولات جادة للعثور على محكات يحدد في ضوءها صدق الاختبار ، ولم يصلح استخدام الرسم كمحك ، كما لم تصلح تقديرات المشرفين المتصلين بأطفال الحضانة ، وكان أنسب المحكات هو تمايز العمر .

اختبارات واطسون وجليزر للتفكير الناقد :

أعد هذا الاختبار جودون واطسون وادوارد جليزر ليزود المفحوص بمجموعة من المواقف والمشكلات التي تتطلب استخدام التفكير الناقد ، ويتكون من ٥ اختبارات فرعية :

١ - الاستنتاج : ويقاس القدرة على التمييز بين الدرجات المختلفة من الصواب والخطأ ، والتوصل إلى استنتاجات معينة على أساس البيانات المعطاة .

٢ - التعرف على الافتراضات : ويقاس القدرة على التعرف على الافتراضات المتضمنة في القضايا المعطاة .

٣ - الاستنباط : ويقاس القدرة على التفكير الاستنباطي من مقدمات معينة .

٤ - التفسير : ويقاس القدرة على الحكم على الشواهد والأدلة ، وللتمييز بين التعميمات التي تبررها الأدلة والتي لا تبررها .

د - تقويم الحجج : وقيس القدرة على التمييز بين الحجج القسوية والضعيفة بالنسبة لمسائل مطروحة أمام المفحوص .

وقد أعد هذا الاختبار في صورته العربية جابر عبد الحميد جابر ربحي حامد هندام (١) .

وقد استخدم الاختبار في عدد من البحوث ومسائل الماجستير والدكتوراه التي تناولت التفكير للنقد وتوافرت عنه بيانات سيكومترية كافية .

بطارية اختبارات الاستعدادات الحسية الحركية للمكفوفين :

أعد هذه البطارية فرج عبد القادر (٢٢) لقياس الاستعدادات الحسية الحركية ، للمكفوفين ، وتتألف من الاختبارات الآتية :

- ١ - اختبار تمييز اللمس
- ٢ - اختبار تمييز الأبعاد .
- ٣ - اختبار تمييز الوزن .
- ٤ - اختبار المهارة اليدوية .
- ٥ - اختبار مهارة الأصابع .
- ٦ - اختبار تآزر اليدين .
- ٧ - اختبار تصنيف الشكل - اللمس .

ولحساب الثبات طبقت الاختبارات على عينة من ٢٥ من المكفوفين وأعيد التطبيق عليهم بفواصل زمنية لا يقل عن شهر وكانت معاملات الارتباط دالة فيما عدا اختبار مهارة الأصبع . وحسب صدق التكوين الفرضي في ضوء ما هو متوقع نظريا عن الفروق بين الجنسين ، واستخدمت معايير العمر العقلي والدرجات الثنائية والدرجات المعيارية .

مقياس التذكر المباشر :

من اعداد وفيه الكمونى (٢٠) فى رسالتها للدكتوراه عام ١٩٧٦ ويتكون من اختبار للتعرف واختبار الاسترجاع بواسطة مثير مساعد ، واختبار الاسترجاع الحر المباشر واختبار سعة الأرقام . وحددت الصدق كما تقول فى ضوء « التقدير الشخصى او الاعتماد على اراء الحكام والخبراء المشهور بكفاءتهم » ، وهى طريقة لا يعتمد بها كثيرا فى حساب الصدق . قد تفيد فى المراحل الأولى لبناء الاختبار ، انما لا تعد طريقة كافية تصل بالاختبار الى درجة من الصلاحية يعتمد عليها ، كما ان الطريقة التى شرحتها الباحثة فى حساب الثبات (٢٠ : ٤ - ٧) غير واضحة . وعلى هذا لا نوصى باستخدام هذا الاختبار بسبب ضعف معاله السيكومترية .

اختبار الأشكال المتضمنة :

هذا الاختبار من اعداد وتكن Witkin وزملائه لقياس احد الأساليب المعرفية التى شاعت فى السنوات الأخيرة وهو بعد الاعتماد على المجال فى مقابل الاستقلال عنه .

وقد نقل هذا الاختبار الى العربية انور الشرقاوى وسليمان الخضرى الشيخ (٨) وحسب ثباته بطريقة التجزئة النصفية من أداء عينة من طلاب جامعة الزقازيق (ن = ١٦٥) وبلغ معامل الثبات للبينين ٧٦ ر ، وللبنات ٧٨ ر . كما استخدمت طريقة جتمان فى حساب الثبات وتكاد تكون القيم الاحصائية متطابقة فى الحالتين .

وأعدت للاختبار معايير فى صورة السباعى المعيارى ، وقد أجريت على الاختبارات بحوث كثيرة تعد نتائجها شواهد على صدق التكوين الفرضى له .

مراجع الباب الرابع

- ١ - أحمد زكى صالح : اختبار القدرات العقلية الأولية . مكتبة النهضة المصرية (ب . ت) .
- ١ ١ - اسحق حنا بطرس : مدى التوافق بين اختيار كليات الهندسة والقدرات اللازمة للنجاح منها ، رسالة ماجستير ، كلية التربية جامعة عين شمس ، ١٩٧١ .
- ٢ - أمال أحمد مختار صادق : تقنين اختبارات ونج للذكاء الموسيقى على البيئة المصرية فى : بحوث فى تقنين الاختبارات النفسية (تحرير فؤاد أبو حطب) المجلد الثانى . مكتبة الأنجلو المصرية ، ١٩٧٩ .
- ٣ - ————— : مقاييس جديدة للقدرة الموسيقية ، فى المرجع نفسه .
- ٤ - ————— : اعداد صورة مختصرة من بطارية سيشور ، فى المرجع نفسه .
- ٥ - أمال أحمد مختار صادق : كراسة تعليمات اختبارات سيشور للقدرة الموسيقية . مكتبة الأنجلو المصرية ، ١٩٧٢ .
- ٦ - أمال أحمد مختار صادق : كراسة تعليمات اختبارات القدرات الموسيقية للأطفال ، بنتلى ، مكتبة الأنجلو المصرية ، ١٩٧٢ .
- ٧ - ————— ، اكرام مطر : القيمة التنبؤية لاختبارات القبول للتعليم الموسيقى فى مصر . المجلة الاجتماعية القومية ، المجلد ١٢ ، العدد ٢ ، سبتمبر ١٩٧٦ .
- ٨ - انور محمد الشرقاوى ، سليمان الخضرى الشيخ : اختبار الأشكال المتضمنة . مكتبة الأنجلو المصرية ١٩٨٥ .
- ٩ - جابر عبد الحميد جابر ، يحيى حامد هندام : كراسة تعليمات اختبارات التفكير الناقد . دار النهضة العربية ، ١٩٧٠ .

- ١٠ - سيد محمد خير الله : اختبار القدرة على التفكير الابتكاري (ب.ت)
- ١١ - سيد محمد خير الله ، محمود عبد الحليم منسى : اختبار النسيير الابتكاري للأطفال (ب.ت) .
- ١٢ - السيد محمد خيرى . علم النفس الصناعى : دار النهضة العربية ١٩٦٧ .
- ١٣ - السيد محمد خيرى : بطارية اختبارات استعدادات حرف المعادن مصلحة الكفاية الانتاجية والتدريب المهني (ب.ت) .
- ١٤ - عبد السلام عبد الغفار : اختبارات القدرة على التفكير الابتكاري . مطبعة البيان العربى ، ١٩٦٥ .
- ١٥ - عبد الله سليمان ، فؤاد أبو حطب : كراسة تعليمات اختبارات تورنس للتفكير الابتكاري . مكتبة الأنجلو المصرية ، ١٩٧١ .
- ١٦ - عبد الله محمود سليمان ، فؤاد أبو حطب : تقنين اختبارات تورنس للتفكير الابتكاري على البيئة المصرية : اختبارات الأشكال (الصورة ا) . فى بحوث فى تقنين الاختبارات النفسية ، (تحرير فؤاد أبو حطب) ، المجلد الأول ، مكتبة الأنجلو المصرية ، ١٩٧٧ .
- ١٧ - عبد الوهاب محمد كامل : اختبار هيكى نيراسكا لقياس الاستعداد للتعلم ، طنطا : المكتبة الفرعية الحديثة (ب.ت) .
- ١٨ - على حسين بدارى ، أنور رياض عبد الرحيم : الاختبارات المعرفية مرجعية العوامل (غير منشور)
- ١٩ - فؤاد أبو حطب ، القدرات العقلية . مكتبة الأنجلو المصرية (ط ٤) ، ١٩٨٣ .

- ٢٠ - ——— ، التفضيل الفنى وسمات الشخصية ، المجلة الاجتماعية القومية ، يناير ١٩٧٢ .
- ٢١ - ——— ، آمال صادق : كراسة تعليمات اختبار مينيسوتا للوحات الأشكال (غير منشور) .
- ١٢١ - فتحى الزيات : تقنين بطارية اختبارات هفلانجان على طلاب كلية الطب . رسالة ماجستير ، كلية التربية جامعة المنصورة ، ١٩٧٨ .
- ٢٢ - فرج عبد القادر طه : اختبارات الاستعدادات الحسية الحركية للمكفوفين . مطبعة دار التأليف ، ١٩٧٤ .
- ٢٣ - محمد ثابت على الدين : اختبارات تورنس للتفكير الابتكارى عند الأطفال باستخدام الحركات والافعال . مكتبة الأنجلو المصرية ، ١٩٨٥ .
- ٢٤ - محمد عبد السلام احمد . القدرة الكتابية واختباران لقياسها : لجنة التأليف والترجمة والنشر ، ١٩٦٠ .
- ٢٥ - محمد عماد الدين اسماعيل . قياس الابداع الفنى وعلاقته بالتربية الفنية ، النهضة المصرية ، ١٩٥٩ .
- ٢٦ - محمد عماد الدين اسماعيل ، سيد عبد الحميد مرسى : كراسة تعليمات اختبارات المهن الكتابية . مكتبة النهضة المصرية (ب)ت
- ٢٧ - محمود السيد ابو النيل : الحد الأدنى اللازم لاداء الاختبارات النفسية للمهنيين . المركز القومى للبحوث الاجتماعية والجنائية ، ١٩٧٥ .
- ٢٨ - محمود عبد القادر : تقنين بطارية الاستعدادات العامة . المركز القومى للبحوث الاجتماعية والجنائية ، (١٩٧٠) .
- ١١٨ - مصطفى خليل الشرقاوى : دراسة لأنماط القدرات اللازمة للتكيف والنجاح فى التدريب على بعض المجموعات المهنية وتطبيقاتها فى مجال

التربية والتوجيه المهني • رسالة دكتوراه ، كلية التربية جامعة عين
شمس ، ١٩٧٢ • أدلة بطارية اختبارات الاستعدادات الفارقة •

٢٩ - هدى برادة ، فاروق صادق : بطارية اختبارات القدرات النفسية
اللغوية وتقنياتها على عينة مصرية في كتاب : بحوث في تقنين
الاختبارات النفسية (تحرير فؤاد أبو حطب) ، المجلد الثاني ، مكتبة
... الأنجلو المصرية ، ١٩٧٩ •

٣٠ - وفية الكمونى : مقياس التذكر المباشر ومتغير المدى • مطبعة النهضة
الجديدة ١٩٧٦ •

31. Anastasi, A. Psychological Testing. Macmillan, 1982.
32. Bennett, G.K., et al. Test of Mechanical Comprehension, New York, Psychol. Corporation, 1940-1954.
33. Anastasi, A. Drake, J. An empirical comparison of certain techniques for estimating the reliability of speeded tests. Educ. psycho. Measment, 1954, 14, 529-540.
34. Cronbach, L.J., Essentials of Psychological testing, Harper, 1984.
35. Flanagan, J. C. Flanagan Aptitude classification Tests. Technical Report, Chicago, SRA, 1959.
36. Fleishman, E.A. Dimentional analysis of movement reactions, J. Exp. Psychol. 1958, 17, 522-532.
37. Freeman, F.S. Theory and Practice of Psychological testing. (3rd ed.), Holt, Rinehart, Winston, 1962.
38. Guilford, J.P. The nature of human intelligence, McGraw-Hill, 1967.
39. Likert, R., Quasha, W.H. Revised Minnesota Paper Form-bord. Psychol. corporation, 1941-1984.

40. Mellenbruch, P.L. mechanical motivation test. Psycho. Affiliates, 1957.
41. Melton, H.W. (ed.) Apparatus Tests. Washington, D.C. Government printing offices, 1947.
42. Mourad S.A. & Moustafa M.M. The Development of an Arabic Aptitude test. A.U.C., Egypt.
43. Nunnally, J.C., (Jr.) Introduction to psychological Measurement. McGraw-Hill, 1970.
44. Richardson, B. et al. SRA Mechanical Aptitudes, Chicago; Science Research Associate, 1949-1950.
45. Super D.E. (ed.) The use of multifactor tests in guidance. Washington, Amer. Pers and Guidance Assoc., 1959.
46. Super, D.E., crites, J.O. Appraising vocational fitness by means of psychological tests. Harper, 1962.

الباب الخامس

اختبارات التحصيل

الفصل الحادى عشر

بناء الاختبار التحصيلى

يقصد بالاختبار التحصيلى achievement test الأداة التى تستخدم فى قياس المعرفة والفهم والمهارة فى مادة دراسية أو تدريبية معينة أو مجموعة من المواد .

وهذه الاختبارات هى أكثر الأدوات شيوعا فى التقويم التربوى ، وهى فى بلادنا هى الوسيلة الوحيدة التى تستخدم فى توجيه التلاميذ وانتقائهم ، كما أنها فى أغلب الأحوال ما يمكن للمعلم أن يصنعه فى ممارسته اليومية للتدريس . ولذلك فإنها تحتل منزلة خاصة بين أدوات التقويم النفسى يجعل من المهم أفراد هذا الفصل لها بحيث يمكن تناولها بشئ من التفصيل حتى يمكن للمعلم ، والسلطات التربوية عامة ، أن تستفيد من المبادئ التى سوف نتناولها هنا فى أعداد هذه الاختبارات .

تحديد الأهداف التعليمية :

أشرنا فى الفصل الأول من هذا الكتاب الى العلاقة بين التقويم النفسى والأهداف التربوية educational objectives . والواقع أن هذه الأهداف تتفاوت من العبارات الشاملة التى تدل على الأغراض العامة والقيم والفلسفات التربوية الى العبارات المحددة التى تبين أساليب أداء نوعية أو خاصة . ويشار الى النوع الثانى فى علم النفس التربوى المعاصر بالأهداف التعليمية instructional objectives ، وإذا كانت الأهداف التربوية يعوزها الوضوح فى أغلب الأحوال ، فإن الأهداف التعليمية أكثر تحديدا لأنها تتناول النواتج النهائية لعملية التعلم والتعليم فى صورة أداء يمكن ملاحظته ، وبالتالي فهى على درجة أكبر من الوضوح .

ويرى كثير من علماء النفس التربويين المعاصرين (١٠) أنه من

الواجب صياغة الأهداف التعليمية فى عبارات صريحة explicit
تدل على الأداء النهائى terminal performance لثلاثة أسباب هى :

١ - يؤدى تحديد الأهداف التعليمية تحديدا سلوكيا بهذا الشكل الى توجيه المعلم فى تخطيط عملية التعليم فيمكنه ان يحدد الخطوات التى يجب ان يسير فيها التلميذ للوصول الى الهدف ، كما يمكن ان ينهى بأعباء جميع الاستجابات التى يجب ان تصدر عن التلميذ للوصول الى الاستجابات الهامة النهائية .

ومن المؤكد ان اساليب التدريس تعتمد على الأهداف التعليمية والمدخلات السلوكية ، ومع ذلك فان من نتائج تخطيط المعلم هذا وبخه قد يكون تعديل الأهداف بتغيير مقدار الزمن مثلا او المستوى المتوقع للالتقان او الموضوعات التى يجب تناولها فى التدريس . وبالطبع لا يمكن لهذه التغييرات ان تحدث ما لم تتم صياغة الأهداف بصورة صريحة . وفى نفس الوقت لا يستطيع المعلم ان يوائم بين اساليبه فى التعليم والفروق فى المدخلات السلوكية ما لم يحدد بشكل دقيق ما يستطيع التلميذ اداؤه قبل التعليم وما يتوقع منه اداؤه بعد التدريس ، ومعنى ذلك ان اهم اسباب وجوب صياغة الأهداف بشكل صريح ما يهيؤه ذلك للمعلم من ارشاد فى تخطيط اساليبه فى التدريس .

٢ - يفيد تحديد الأهداف التعليمية فى عبارات صريحة فى تقويم الأداء ومن المعروف ان الاهتمام بالصياغة الدقيقة للأهداف التعليمية بدأه فى الأصل العلماء المهتمون ببناء الاختبارات وتقويم المناهج المدرسية ، فقد اكتشفوا ان استخدام العبارات الغامضة فى التعبير عن الأهداف يجعل عملية بناء الاختبار ومفرداته عملية صعبة ان لم تكن مستحيلة .

٣ - تفيد ترجمة الأهداف التعليمية فى صورة أداء نهائى فى توجيه انتباه التلميذ وجهوده .

وقد اثبت ميجر ومكان (١٩) ان التعلم يمكن ان يحدث حين لا يفعل

المعلم أكثر من إعطاء التلاميذ قائمة بالأهداف التعليمية الواضحة يقوم التلاميذ بالاهتمام بها في تعلمهم الذاتي .

وصف العمل التربوي :

'يتطلب تحديد الأهداف التعليمية وصفا دقيقا للعمل التربوي كما يتطلب تحليلا تفصيليا لهذا العمل . وطريقتا وصف العمل task description وتحليل العمل task analysis نشأتا أول الأمر في ميدان علم النفس الصناعي ثم انتقلتا في السنوات الأخيرة إلى ميدان علم النفس التربوي .

ومن أهم أساليب وصف العمل التربوي طريقة ميجر Mager وطريقة ميللر Miller ، وتتطلب طريقة ميجر (٧ ، ١١) في وصف العمل - أو بعبارة أخرى تحديد الأهداف التعليمية - خطوات ثلاثة هي :

- ١ - تحديد الأداء النهائي الذي يسعى التعليم أو التدريس إلى إنتاجه .
- ٢ - وصف الظروف والأحوال الهامة التي تتوقع حدوث السلوك فيها .
- ٣ - وصف مستوى الجودة الذي يجب أن يصل إليه أداء التلميذ .

فمثلا قد يكون الهدف التعليمي الذي نسعى إلى تحقيقه هو أن يستطيع التلميذ تحديد العوامل التي أدت إلى حدوث الحرب العالمية الأولى ، وهذا هدف مترجم في أداء يمكن ملاحظته ملاحظة صريحة وعامة في صورة أداء نهائي يسعى التعليم إلى إنتاجه (الخطوة الأولى) ، ثم يمكن أن نعطي للتلميذ قائمة بالعوامل التي تؤدي إلى الأحداث التاريخية الهامة ، وتكون هذه القائمة وصفا لما يسميه ميجر العوامل الهامة التي يحدث في ظلها السلوك النهائي (الخطوة الثانية) ، وأخيرا تتحدد المستويات المطلوبة (الخطوة الثالثة) . والواقع أن تحديد المستويات هذا ليس من الأمور القاطعة ويتراوح بين الاتقان الكامل وبين أقل مستويات الاتقان .

وتشير طريقة ميللر (١٠) في وصف العمل إلى وجود ثلاث خطوات أيضا هي :

١ - ظروف الاستثارة stimulus condition أو ما يسمى به indication .

٢ - التنشيط activation أو الاستجابة التي يجب أن تصدر عن التلميذ .

٣ - التغذية الراجعة Feedback أو تحديد مستوى ملاءمة الاستجابة .

والواقع أن هذه الخطوات لا تختلف كثيرا عن تلك التي أشار إليها ميجر ، فظروف الاستثارة هي ما يشير إليه ميجر بالعوامل أو الظروف أو الأحوال الهامة ، والتنشيط عند ميللر يقابل ما يسميه ميجر بالأداء النهائي ، ويتشابه مفهوم التغذية الراجعة عند الأول مع مفهوم مستوى الأداء عند الثاني .

تحليل العمل التربوي :

بعد ما يقوم المعلم بوصف الأعمال التي يتوقع من تلاميذه أداءها يكرن عليه تحليل هذه الأعمال وتصنيفها . والغرض الرئيسي من تحليل العمل هو مساعدة المعلم على تحديد أنسب ظروف التعلم ، والتي تختلف من نوع من أنواع التعلم لآخر . فتعلم تكوين المفاهيم أو حل المشكلات أو الابتكار يتطلب ظروفا تختلف عن اكتساب المهارات الحركية أو حفظ قائمة من الكلمات . وهذا يعني بالطبع أن مسئولية المعلم أكبر من مجرد وصف العمل التربوي أو صياغة الأهداف في عبارات سلوكية صريحة . فهي تشمل أيضا تصنيف هذه الأساليب السلوكية المتضمنة في كل هدف من هذه الأهداف ثم تحديد أنسب شروط التعلم لكل منها .

ويوجد في الوقت الحاضر تصنيفان أساسيان للسلوك لهما أهميتهما التربوية هما تصنيف جانيم Gagné وتصنيف بلوم Bloom . ويشمل تصنيف جانيم (١٤) ثمانية أنواع من التعلم تتدرج من البسيط الى المركب هي:

١ - التعلم الاشاري Signal learning أو ما يسمى في سيكولوجية

التعلم بالاشتراط الكلاسيكى وفيه يتم تعلم استجابة شرطية لمثير او
اشارة معينة .

٢ - تعلم المثير - الاستجابة او ما يسمى بالاشتراط الاجرائى .

٣ - التسلسل chaining او ما يسعى اكتساب المهارات الحركية
(غير اللفظية) .

٤ - الارتباط اللفوى Verbal association .

٥ - التمييز المتعدد Multiple discrimination

٦ - تعلم المفاهيم Concept learning .

٧ - تعلم المبادئ Principle learning .

٨ - حل المشكلة Problem solving او التفكير .

اما التصنيف الذى وضعه بلوم Bloom ومساعدوه للميدان العقلى
المعرفى (٩) فيشتمل على فئتين كبيرتين هما المعرفة Knowledge
والقدرات والمهارات العقلية بحيث نحصل فى النهاية على ستة فئات للاهداف
التربوية هي :

١ - المعرفة Knowledge .

٢ - الفهم Comprehension .

٣ - التطبيق Application .

٤ - التحليل Analysis .

٥ - التركيب Synthesis .

٦ - التقويم Evaluation .

وصف المدخلات السلوكية :

تتطلب عملية التعليم وصفا للمدخلات السلوكية لدى التلاميذ حتى
يمكن للمعلم ان يخطط لتنفيذ عملية التدريس ، بل وحتى يمكن للمسلطات
التربوية ان تخطط للعملية التربوية عامة ويتطلب هذا الوصف تحديدا دقيقا

الموضع الراهن لمعارف التلميذ ومعلوماته ومهارته وطرق تفكيره وقدراته واستعداداته وميوله ودوافعه وخصائص شخصيته ، قبل أن نسمى الى اكسابه النماذج السلوكية الجديدة من خلال التعلم والتدريس . وهذه البيانات يجب استخدامها لأغراض الوصف وليس لأغراض التفسير ، لأن هذه المفاهيم تصف في الواقع المستوى السلوكي للفرد في لحظة معينة ، ولا تستطيع الاختبارات والمقاييس التي نستخدمها في قياس هذه الظواهر أن تفسر لنا أداء الفرد فيها ، فلا يمكن مثلا أن نعزو ضعف الأداء في اختبار للذكاء الى « ضعف الذكاء » ، لأن ذلك يصبح نوعا من تحصيل الحاصل ولا يزيد فهمنا للسلوك الانساني .

والواقع أن وصف المدخلات السلوكية ضروري لأي برنامج تربوي . فعملية التعليم يجب أن تبدأ مما لدى التلاميذ ويجب أن تستمر الى حيث يستطيعون من أداء نهائي تحت شروط معينة ومستوى معين . كما أن تحديد اهدافنا يمكننا من أن نقرر ما سوف نعلمه وما نفترض أنه لدى التلاميذ بالفعل ويمكن للتفاعل بين الأهداف والمدخلات السلوكية أن يؤدي الى تغيير الأهداف اذا ثبت أنها غير واقعية ، كما تتفاعل المدخلات السلوكية مع وسائل التدريس واساليبها ، فنحن لا نبدأ تدريسنا لتحقيق هدف تربوي معين بمواد تعليمية جديدة تماما وانما قد نستدعي الاستجابات التي سبق تعلمها وخاصة تلك التي تتناسب مع موقف التعلم الجديد .

ثم ان لكل نوع من التعلم مدخلاته السلوكية فتتمة سلوك حل المشكلة تتطلب معرفة بمدخلات سلوكية تختلف عما يتطلبه تنمية احسدى المهارات الحركية (٦) .

العلاقة بين الأهداف التعليمية واساليب التقويم التربوي :

يلخص جرونلاند (١٥) العلاقة بين الأهداف التربوية والأهداف التعليمية واساليب التقويم التربوي على النحو الآتي :

الأهداف التربوية العامة » الأهداف التي توجه التدريس توجيهها يوميا،



نواتج التعلم الخاصة أو النوعية » الأهداف التعليمية ،

» سلوك التلاميذ وأداؤهم الظاهر الذي يعد دليلا عن تحقيق هذه الأهداف،



أساليب التقويم

(الأساليب التي يحصل بها على عينة من سلوك التلاميذ كما توصف في

نواتج التعليم النوعية) •

ويتطلب هذا ربط أساليب التقويم مباشرة بنواتج التعليم التي نقومها وبهذا نتيقن من أننا نقوم التلميذ نحو الأهداف التي أخذناها كإطار لعملية التدريس • وعملية ربط أساليب التقويم بنواتج التعلم النوعية هي في أساسها عملية تحليل منطقي وحكم • هذه العملية يمكن تيسيرها باستخدام خطة ما من خطط التقويم المنظمة ، هذه الخطة يجب أن تشمل على الأقل قائمة بنواتج التعلم المرغوب فيها والأساليب التي تستخدم في تقويم التقدم نحوها • واليك خطة على سبيل المثال في ميدان علم الأحياء للصف الأول الثانوي تشمل الأهداف التعليمية ونواتج التعلم النوعية وأساليب التقويم الملائمة •

١ - يعرف التلميذ المصطلحات الشائعة المستخدمة في علم الأحياء حيث:
١ - يعرف المصطلحات الشائعة في علم أساس المعنى ١ - ٢ - اختبار موضوعي
الإجابة القصيرة

٢ - ١ يميز بين المصطلحات الشائعة على أساس البنى ١ - ٢ اختبار موضوعي
١ - ٢ يتعرف على معاني المصطلحات الشائعة حين تستخدم في نص معين
١ - اختبار موضوعي •

٢ - يظهر التلميذ القدرة على التفكير الناقد حينما :

٢ - ١ يميز بين الحقائق والآراء
١ - ٢ اختبار موضوعي

٢ - ٢ يشق استنتاجات صحيحة من المعطيات ٢ - ٢ اختبار مقال من نوع
الاجابة القصيرة .

٢ - ٢ يتعرف على الافتراضات وراء الاستنتاجات ٢ - ٢ اختبار موضوعي
٢ - ٤ يتعرف على نواحي القصور في المعطيات ٢ - ٤ اختبار موضوعي

٢ - يؤدي التلميذ الأعمال الأساسية للتشريع بمهارة حينما :

٢ - يضع العينة في الوضع الصحيح للتشريع ٢ - ١ قائمة ملاحظة أو
مقياس تقدير

٢ - ٢ يقطع بمهارة دون افساد العينة التي يدرسها ٢ - ٢ قائمة ملاحظة
أو مقياس تقدير .

٢ - ٣ يقسم اجزاء البنية في العينة دون افسادها ٢ - ٢ قائمة ملاحظة
أو مقياس تقدير .

٢ - ٤ ينتهي من التشريع في الوقت المحدد ٢ - ٤ قائمة ملاحظة أو
مقياس تقدير .

٤ - يظهر التلميذ القدرة على تحديد مواضع المعلومات البيولوجية
حينما :

٤ - ١ يستخدم كتالوج المكتبة في تحديد المراجع ٤ - ١ تقرير بحث أو ملاحظة
٤ - ٢ يعرف المصادر المشتركة للمعلومات البيولوجية ٤ - ٢ اختبار
موضوعي .

٤ - ٣ يستخدم جداول المحتوى والفهرس عند

البحث عن المعلومات في الكتب ٤ - ٣ ملاحظة

٤ - ٤ يحدد مدى ملاءمة المعلومات لمشكلة محددة ٤ - ٤ تقرير بحث أو ملاحظة

٥ . يكون لدى التلميذ اتجاه علمي نحو الظواهر البيولوجية حينما :

٥ - ١ يؤجل الحكم حتى تتوافر جميع الحقائق أو معظمها .

٥ - ١ تقرير قصصي ، اختبار موضوعي

- ٥ - ٢ يبحث عن علاقات السبب والأثر فى البيانات البيولوجية
- ٥ - ٢ تقرير قصصى • اختبار موضوعى
- ٥ - ٢ يظهر رغبة فى التفاسير الجديدة للبيانات البيولوجية
- ٥ - ٢ تقرير قصصى • اختبار موضوعى
- ٥ - ٤ يفسر البيانات البيولوجية متحررا من التحيز
- ٥ - ٤ تقرير قصصى • اختبار موضوعى
- ٥ - ٥ يظهر ثقة فى البيانات البيولوجية التى يحصل عليها بالأساليب العلمية
- ٥ - ٥ تقرير قصصى • اختبار موضوعى

من هذا الجدول نجد ما يأتى :

١ - نواتج التعلم النوعية كما تصاغ فى صورة سلوك للتلاميذ عديدة ومتنوعة بحيث لا توجد أداة تقويم واحدة تعطى دليلا مباشرا على تحصيل هذه الأهداف التربوية • ولذلك نحتاج الى جانب الاختبارات الموضوعية الى قوائم الملاحظة والتقارير القصصية واختبارات المقال

٢ - تظهر أهمية الصياغة الواضحة للأهداف ونواتج التعلم فى انتقاء أداة التقويم • وفى الواقع حينما تصاغ نواتج التعلم فى ضوء سلوك التلميذ فإنها لا تدل فقط على ماذا نقوم ولكنها توحى أيضا بكيف نقوم •

مثال : (١ - ١) • يحدد المصطلحات الشائعة ، يوضح لنا أسلوب التقويم الذى يجب أن يستخدمه • فهى تدل على أن التلميذ يجب أن يعطى التعاريف بنفسه وعلى ذلك فإن اختبار المقال من نوع الإجابة القصيرة حيث يعطى للتلميذ عدة مصطلحات ويطلب منه تعريفها هو الأسلوب الأكثر ملاءمة لتقويم هذا الهدف • أما الاختبار الموضوعى مثلاً من نوع الاختيار من متعدد حيث يطلب من التلميذ أن يتعرفوا فقط على التعريف لا تتناسب مع تقويم هذا الناتج التعلمى فى صيغته هذه ، والا فإن للصيغة تصبح • يتعرف التلميذ على معنى المصطلحات الشائعة • وفى هذه الحالة يمكن استخدام الاختبار الموضوعى • ويتوقف هذا على غرضنا من التقويم • فإذا كنا نريد أن نتأكد من أن التلميذ يعرفون المصطلحات عن طريق تحديدهم وتعريفهم لها فإن الإجراء الوحيد الملائم هو أن يطلب من التلميذ أن يعرفوا •

مصطلحاتهم • والقدرة على التعرف الصحيح لا يمكن اعتبارها دليلا على قدرة التلاميذ على اعطاء التعريف الصحيح ما لم يكن هناك دليل من البحث العلمى يؤكد أن نوعى السلوك مرتبطان ارتباطا عاليا • وحيث انه لا يوجد مثل هذا الدليل بعد فان الاجراء الأحوط هو اختيار أسلوب التقويم الذى يرتبط مباشرة بنتائج التعلم النوعى (الهدف التعليمى) •

تحليل محتوى المادة الدراسية :

أشرنا الى أن الاختبار التحصيلى يرتبط بمادة دراسية معينة نم تدريسها بالفعل ، ومعنى ذلك أنه لا يوجد ما يبرر اعداد اختبارات من هذا القبيل تتناول مواد لم يتم تدريسها أو لم تكن محددة كمحتوى لعملية التعلم • ولذلك لا بد فى اعداد الاختبار التحصيلى من تمثيل كل ما حظى باهتمامنا أثناء عملية التدريس • ومن الطرق الملائمة فى تحديد الأصل الاحصائى المناسب الذى نشق منه عينة السلوك التى تسمى الاختبار التحصيلى أن نعد قائمة بالموضوعات التى تتضمنها هذه المادة بأكبر قدر من التفصيل وفى أبسط صورة ممكنة • ويمكن للباحث فى هذا الصدد أن يستفيد بمنهج البحث المعروف باسم « تحليل المحتوى » Content analysis (٥) •

تحديد الوزن النسبى للإهداف التعليمية ومحتوى المادة الدراسية :

بعد تحليل المحتوى يقوم الباحث بتحديد الوزن النسبى لكل موضوع (أو وحدة) من موضوعاته ، ومن المحكات المستخدمة فى تحديد وزن كل موضوع من موضوعات المحتوى هو الزمن المخصص لتدريس كل منها • إلا أنه فى بعض الأحيان لا يكون الوقت المستغرق فى تدريس موضوع معين دلالة على أهميته ، ولذلك فان الأوزان النسبية للموضوعات لا بد أن تعكس اهتمام المعلم به • وبالمثل لا بد من أن يحدد المعلم الأوزان النسبية للإهداف التعليمية لمادته • فإذا كان يؤكد فى تدريسه وصول التلاميذ الى التعميمات وتطبيقها مثلا فيجب أن يعطى هذا الهدف وزنا أكبر • وبعبارة موجزة لا بد أن تعطى لموضوعات المحتوى وللأهداف التعليمية أوزان نسبية تعكس الاهتمام الذى تحظى به فى عملية التدريس •

وفى تحديد هذه الأوزان قد يعتمد المعلم على خبرته فى تدريس المادة وتحليله لمحتواها فيضع قائمة بأهم الموضوعات ثم يصنفها حسب الأهداف التى يرى أن يقيسها الاختبار . إلا أنه فى كثير من الأحيان لا يكتفى الباحث بالاعتماد على خبرته الشخصية وحدها ويلجأ الى بعض الخبراء لاستشارتهم فى تحديد هذه الأوزان النسبية .

اعداد جدول المواصفات :

على أساس المعلومات التى يتم جمعها يتم تحديد مواصفات الاختبار test specifications والتى تتضمن الموضوعات التى يجب أن يشملها الاختبار ونواتج التعلم التى يجب اختبارها (أى الأهداف التعليمية) ، والأهمية النسبية (أو الوزن النسبى) للموضوعات والأهداف . وعلى هذا الأساس يتم تحديد عدد الأسئلة التى يتم اعدادها مرتبطة بكل هدف فى كل موضوع . والطريقة الأكثر سهولة فى تحديد مثل هذه المواصفات هى اعداد جدول المواصفات table of specifications كما تناولنا فى الفصل الخامس .

تحديد نوع الأسئلة :

أشهر أنواع الأسئلة التى تستخدم فى ميدان الاختبارات التحصيلية أسئلة المقال والأسئلة الموضوعية ، وأسئلة حل المشكلات (أو الأسئلة التفسيرية) . ونعرض فيما يلى عرضاً موجزاً لأنواع الأسئلة الشائعة فى التقويم التربوى الحديث .

أسئلة الاختيار من متعدد .

وتعتبر أكثر الأسئلة الموضوعية شيوعاً ، وتقيس بكفاءة شديدة النواتج البسيطة للتعلم ، بل قد تصلح لقياس بعض النواتج التعليمية المعقدة . ويتكون السؤال من مشكلة (قد تصاغ فى صورة سؤال مباشر أو عبارة ناقصة) تسمى الجذر Stem ، وقائمة من الحلول المقترحة (قد تشمل كلمات أو أعداداً أو رموزاً أو عبارات) وتسمى البدائل الاختيارية alternatives . ويطلب من المفحوص قراءة جذر السؤال وقائمة البدائل وانتقاء البديل

الصحيح أو الأفضل . والبديل الصحيح فى كل سؤال يسمى ببساطة الاجابة
answer بينما تسمى البدائل الاختيارية المسؤلات (أو المشتتات)
distractors ، وهى تسمى كذلك من واقع وظيفتها المقصودة . أى
تحويل هؤلاء الذين فى شك من الاجابة الصحيحة .

ويستخدم هذا النوع من الأسئلة فى قياس كثير من الاهداف التعليمية
وخاصة نواتج التعلم فى ميدان التذكر واكتساب المعلومات ، ومن ذلك
معرفة المصطلحات ومعرفة الحقائق النوعية ، ومعرفة المبادئ والقوانين ،
ومعرفة الطرق والاجراءات ، كما تستخدم فى قياس الفهم مثل القدرة على
تطبيق الحقائق والمبادئ ، والقدرة على تفسير علاقات السبب والاثـر ،
والقدرة على تبرير الطرق والاجراءات .

ويتميز هذا النوع بمرونته الشديدة ، وتحرره من نقائص الأنواع
الأخرى من الأسئلة ، ومن مشكلة اسلوب الاستجابة response set
بل ان صياغة البدائل الاختيارية قد يؤدى استخدامها فى أغراض التشخيص .
ومع ذلك فان لأسئلة الاختيار من متعدد مشكلاتها وأهمها انها تقتصر على
المستوى اللفظى للاداء (وهى فى هذا تشترك مع غيرها من أسئلة الورقة
والقلم) ، كما انها لا تصلح لقياس بعض مهارات حل المشكلة فى الرياضيات
والعلوم ، والقدرة على تنظيم الأفكار (وهى فى هذا تشترك مع غيرها من
أسئلة التعرف على الاجابة) . ثم ان لأسئلة الاختيار من متعدد نقيصة
خاصة بها لا تشترك فيها مع غيرها ، وهى صعوبة ايجاد عدد كاف من
الحولات المحبذة وغير الصحيحة فى وقت واحد . الا ان هذه الصعوبة
تتناقص مع زيادة الخبرة فى بناء مثل هذه الأسئلة .

أسئلة التفسير :

يتطلب هذا النوع من الأسئلة ما هو أكثر من الحفظ واكتساب المعلومات
الجزئية ، أى انه يستخدم فى قياس نواتج التعلم المرتبطة بالعمليات العقلية
العليا كالفهم ومهارات التفكير وقدرات حل المشكلة . ويتكون السؤال
التفسيري interpretative من سلسلة من الأسئلة الموضوعية تعتمد
على مجموعة مشتركة من البيانات الأولية (المعطيات) . وقد تكون هذه

المعطيات فى صورة مواد مكتوبة أو جداول أو رسوم أو أشكال أو خرائط أو صور • وقد تتخذ الأسئلة المرتبطة صوراً مختلفة ولكنها فى الغالب تأخذ صورة الاختيار من متعدد •

أسئلة المزاوجة :

تتكون أسئلة المزاوجة matching التقليدية من عمودين متوازنين يحتوى كل منهما على مجموعة من العبارات أو الرموز أو الكلمات • وتسمى المفردات فى العمود الذى تنشده المزاوجة بالمقدمات premises ، وتسمى المفردات التى نختار منها بالاستجابات responses .

وعادة ما تكون أسس مزاوجة الاستجابات للمقدمات واضحة فى ذاتها ، وفى بعض الأحيان تتطلب توضيحاً فى التعليمات • ويستخدم هذا النوع فى قياس الحقائق التى تستند الى التداعى البسيط مثل الشخصيات والانجازات ، التواريخ والأحداث التاريخية ، المصطلحات والتعاريف ، القواعد والأمثلة ، الرموز والمفاهيم ، المؤلفون ومؤلفاتهم ، الكلمات الأجنبية ومقابلاتها بالعربية الآلات واستخداماتها ، النبات وتصنيفاته ، الحيوان وتصنيفه ، الأشياء وأسمائها ، الأعضاء ووظائفها • وتتميز بأنها تقيس بدقة وسرعة مقداراً كبيراً من الحقائق المرتبطة فى زمن قصير نسبياً ، بالإضافة الى أنها سهلة الصنع • ومع ذلك فإن لهذا النوع مشكلاته وأهمها أنه يقتصر على قياس الحقائق التى تعتمد على التذكر والتداعى البسيط • وكذلك يصعب على الباحث إيجاد مادة متجانسة لها أهميتها فى المحتوى الدراسى يمكن أن تصاغ منها قائمة المقدمات •

أسئلة الإجابة القصيرة (الاستدعاء) :

يتطلب هذا النوع من الأسئلة الذى يعرف فى بعض المؤلفات باسم أسئلة التكملة - أن ينتج المفحوص استجابته - وليس مجرد التعرف عليها كما هو الحال فى الأنواع السابقة • ويمكن القول أن هذا النوع قد يتطلب إجابة قصيرة إذا عرضت المشكلة فى صورة سؤال مباشر ، أو تكملة إذا عرضت فى صورة عبارة ناقصة • وتوجد أنواع أخرى من هذا النوع منها : أسئلة

اعداد القوائم Listing والتي تسمى احيانا اختبارات المقال ذات التعليمات المحددة (مثل اذكر ٣ اسباب لقيام الحرب العالمية الأولى) واسئلة التماثل أو التناسب analogy . واسئلة المشكلات ، واسئلة التعيين identification (كُن يطلب من التلميذ تعيين اجزاء رسم نبات أو حيوان) . ويستخدم هذا النوع في قياس النواتج البسيطة نسبيا للتعلم مثل معرفة المصطلحات ، ومعرفة الحقائق النوعية ، ومعرفة المبادئ ، ومعرفة الطرق والاجراءات ، والتفسيرات البسيطة ، والقدرة على حل المسائل في الرياضيات والعلوم . وهذا النوع ربما بسبب بساطة نتائج التعلم التي تقيسها (الاستدعاء في اغلب الأحوال) ، بالاضافة الى تحرره من اثر التخمين السائد في اختبارات التعرف بوجه عام . الا ان اكثر مشكلاته شيوعاً هي صعوبة تصحيحه .

اسئلة الصواب والخطا (أو اسئلة البديلين alternative response) :

وتتطلب اختيار اجابة واحدة من اجابتين كالحكم على العبارة بالصواب أو الخطأ ، أو الاجابة على السؤال بنعم أو لا ، أو الحكم على العبارة بأنها تدل على رأى أو حقيقة ، و تقدير عبارة بالموافقة و المعارضة . ويستخدم هذا النوع في قياس نتائج التعلم التمييزي البسيط ، ومع ذلك فيمكن أن تفيد في قياس الفهم (كالتعرف على علاقات السبب والاثـر) . وقد يكون اكثر مشكلات هذا النوع من الاسئلة خطورة هو تأثرها البالغ بالتخمين .

اسئلة الترتيب :

وفيها يقوم الفحوص باعادة ترتيب خطوات أو أحداث أو اجراءات أو تواريخ في تسلسل طبيعي أو منطقي .

اسئلة المقال :

رغم الاستخدام الواسع للاسئلة الموضوعية التي اشرفنا اليها لازالت توجد نتائج هامة للتعلم لا يصلح لها الا سؤال ، المقال ، ومنها القدرة على عرض وتنظيم وتكامل الأفكار ، والقدرة على التعبير الكتابي ،

والقدرة على اعطاء التفسيرات والتطبيقات للمعلومات لا مجرد استدعاؤها
أو التعرف عليها .

ومن اهم خصائص اسئلة المقال حرية الاستجابة ، وفي هذا جدارها
كمقياس للتحصيل المعقد . وفيه أيضا تكمن صعوبات التصحيح التي تجعل
منها اداة اقل كفاية وكفاءة في قياس الحقائق والمعلومات . ومعنى ذلك أن
اسئلة المقال يجب أن تستخدم حيث لا تصلح الاختبارات الموضوعية .

الاختبارات العملية :

تستخدم هذه الاختبارات كوسائل موضوعية لتقدير الكفاءة التي يؤدي
بها العمل (كالكتابة على الآلة الكاتبة) أو قيادة السيارات . وتنقسم هذه
الاختبارات الى ٣ انواع :

(أ) اختبارات التعرف : وتتطلب من المفحوص التعرف على الخصائص
الأساسية للاداء (كأن تعترف مقطوعة موسيقية ويطلب من المفحوص تحديد
الأخطاء أو النقائص في الاداء أو التعبير) ، أو تحديد الأجزاء التي تتألف
منها إحدى الآلات ، أو اختيار الآلة أو الجهاز المناسب لعمل معين ، أو تحديد
العينات أو تصنيف الأشياء ، أو انتقاء عمل أدبي أو فني ممتاز .

(ب) الاختبارات التي تتضمن مواقف تشبه المواقف الطبيعية : فهي
تهدف الى قياس الأنشطة الأساسية في العمل . وتسمى أحيانا اختبارات
النماذج المصغرة miniature .

(ج) اختبار عينة العمل : وهو عبارة عن محاولة « مضبوطة » أو
« مقننة » ، في الظروف الواقعية للعمل ، وتنقسم هذه الاختبارات الى نوعين
أساسيين : أولهما الاختبارات التي يسهل فيها التمييز بين الصواب والخطأ
في الاداء وبالتالي يمكن تصحيحها بسهولة مثل التصويب ، والاداء العضلي
في التربية البدنية ، والتجميع الميكانيكي ، والكتسابة على الآلة الكاتبة .
والنوع الثاني هو الاختبارات التي تعتمد على حكم المراقبين والفاخصين
لتقويم الاداء واعطاء درجة أو رتبة ، ومن أمثلة ذلك محاولة سامي ابراهيم
(التلويح النفسي)

على (٤) « تقنين » اختبارات العزف على البيانو . ويتطلب هذا النوع استخدام مقاييس التقدير أو قوائم الملاحظة .

الاختبارات الشفوية :

الاختبار الشفوي هو مزيج من اختبار المقال والاختبار العملي ، وله فائدته في دراسة العمليات العقلية التي يستخدمها المفحوص في الاجابة على أسئلة معينة ، ولذلك فهو أداة نافعة في تشخيص الصعوبات . بل انه في بعض الأحوال هو الأسلوب الأوحس في تقويم المفحوص كما هو الحال في قياس قدرات صغار التلاميذ (في الصفوف الأولى من المرحلة الابتدائية) ، وفي قياس بعض نتائج التعلم اللغوي (كالقراءة الجهرية) .

تحديد عدد الأسئلة :

يتحدد عدد الأسئلة التي يتكون منها الاختبار بمقدار الزمن الميسر له . وكثير من الاختبارات التحصيلية يتحدد زمنه بحوالي ٥٠ دقيقة (زمن الحصّة المدرسية العادية) وبعض جداول الامتحانات تحدد فترات أطول تصل الى ٣ ساعات . وعلى وجه العموم كلما طال الزمن وزاد عدد الأسئلة كان الاختبار أكثر ثباتا . ومن النادر من الوجهة العملية أن نعد اختبارا يستغرق أكثر من ثلاث ساعات .

ومن المفيد كما أشرنا أن نعتبر مجموعة الأسئلة التي يكون منها الاختبار عينة من أصل افتراضي لكل الأسئلة الممكنة التي يمكن أن تستخدم في الاختبار . فمدرس الصف الخامس الابتدائي قد يحصل على ١٠٠ كلمة لاختبار المفردات اللغوية وذلك بانتقاء الكلمة الخامسة من قائمة تحتوي على ٥٠٠ كلمة درست أثناء العام الدراسي . وفي هذه الحالة تصبح الـ ٥٠٠ كلمة هي الأصل الذي تؤخذ منه عينة الاختبار ، وهو أصل حقيقي ومحدد المعالم . الا انه في معظم الاختبارات لا نجد مثل هذا التحديد الحاسم للأصل الاحصائي العام للأسئلة ، فمثلا لا يوجد حد أقصى لعدد المشكلات والمسائل المختلفة التي يمكن اعدادها في اختبارات الجبر .

وبالطبع كلما اتسع مدى الموضوع أو القدرات التي يهدف الاختبار

الى قياسها زاد عدد الأصل الاحصائي . وحجم هذا الأصل يمثل الحد الأدنى لحجم عينة الاختبار . ومعنى ذلك أن أسئلة الاختبار لا يمكن أن تكون أكثر عددا من الأصل التي تشتق منه ، إلا أن حجم هذا الأصل لا يصنع حدا أدنى لحجم العينة . فالأصل الذى يتألف من ١٠٠٠ سؤال يمكن أن تكرر العينة المشتقة منه ١٠ أسئلة أو ٥٠ سؤالا أو ٢٠٠ سؤال وكلما زاد حجم الأصل قل تجانسها ، أى أصبح يحتوى ميادين مختلفة وشبه مستقلة من المعرفة والقدرة . وللوصول الى نتائج دقيقة يجب أن تكون عينة الاختبار المشتقة من أصل غير متجانس أكبر نسبيا من تلك التي تشتق من أصل متجانس . ومع ذلك فلا توجد قاعدة عامة تحدد العلاقة من حجم الأصل والعينة .

ويوجد فى الوقت الحاضر اتجاه يؤكد استخدام عدد قليل وكاف من الأسئلة بحيث يعطى معظم المفحوصين وقتا يحارلون فيه جميع الأسئلة بمعدلاتهم الخاصة العادية . ولشيوع هذا الاتجاه ثلاثة أسباب (١٣) .

١ - أن سرعة الاستجابة ليست هدفا أوليا للتعلم فى معظم المقررات الدراسية وبالتالي لا تعد مؤشرا صادقا للتحصيل لأنه فى كثير من ميادين التحصيل لا ترتبط السرعة والسرعة ارتباطا عاليا . وبالطبع فقد تكون اختبارات السرعة أكثر ملاءمة فى بعض الحالات ، إلا أن هذه المواقف هى الاستثناء وليس القاعدة .

٢ - أن قلق الاختبار قد يكون على درجة كبيرة من الحدة - حتى فى الاختبارات غير الموقوتة - حين يتزايد التأكيد على العمل بسرعة .

٣ - أن الاستخدام الكفء للاختبار التحصيلي - الذى يصنعه المعلم - يتطلب أن معظم التلاميذ يستجيبون له كله حتى يستخدم نتائجه فى أغراض التشخيص .

تقويم التحصيل فى القراءة فى المدرسة الابتدائية (*) :

قد تكون القدرة على القراءة أهم القدرات التي تؤكد لها المدرسة

(*) الامثلة التي سوف نعرض لها مأخوذة بكثير من التعديل عن لندفال (١٧) .

الابتدائية . وفى الصفوف الأولى من المرحلة الابتدائية يحتل تعليم القراءة كمهارة أساسية الجزء الأكبر من العمل المدرسى ، ومع انتقال التلميذ الى الصفوف العليا تتفرع دراسته الى مجالات متنوعة ويقل الوقت المخصص لتعليم القراءة ، ومع ذلك فان لها اهميتها طوال المرحلة الابتدائية . وبعدها يتقن التلميذ بعض المهارات الأساسية فى القراءة ، قد تتعدل تعديلات طفيفة اهداف تدريس القراءة بحيث تؤكد التغيرات الناتجة فى المهارات التى نقيسها . ومن ناحية أخرى فانه لا زال من الاهداف الأساسية للتعلم فهم معانى الكلمات وفهم أنجمل والفقرات الطويلة ، وعلى جميعا لها اهميتها فى جميع المستويات ، ومعنى هذا ان أى اقتراحات تتعلق بتقويم اهداف القراءة فى أحد صفوف المرحلة ينطبق على كل المرحلة .

وفيما يلى مثال لاهداف نوعية لتعليم القراءة فى الصف الأول الابتدائى

الغرض العام	الهدف المرحلى (فى نصف العام)	الاهداف النوعية
أولا : أن يستطيع الفرد المتعلم قراءة اللغة القروية ، قراءة فعالة .	١ - أن يستطيع التلميذ قراءة القسم الأول من كتاب القراءة بسرعة وفهم	١ - أن يستطيع الطفل أن ينطق بكلمات مثل مدرسة ، الحديقة ، ورد ، حينما يراها مكتوبة (عدد كلمات القسم الأول من القراءة ٢٥٠) .
	٢ - أن يستطيع الطفل أن يربط كلمات مثل ولد ، طيارة ، بيت وبين الصور التى تعبر عنها .	
	٣ - أن يستطيع الطفل أن يجيب على أسئلة تدور حول مواد جديدة يقرأها من مختارات من كتاب القراءة .	
ب - أن يستطيع الطفل القراءة الجهرية بسهولة مع بعض التعبير	٤ - أن يستطيع الطفل يغير من نبرات الصوت للتمييز بين الجملة الخبرية والجملة الاستفامية .	

ومن الواضح انه لو كان الواجب هو تقويم هذه الاهداف النوعية جميعا فيجب ان تستخدم اساليب مختلفة . فالهدفان ١ ، ٤ يمكن تقويمها ببعض صور الاداء الشفوي . بينما الهدفان ٢ ، ٣ يمكن تقويمهما بهذه الطريقة أو باستخدام بعض المواد المكتوبة . وبسبب الفوائد الكثيرة الخاصة باستخدام الاختبارات الموضوعية في المدرسة الابتدائية فان على المعلم أن يختار لقياس تحصيل الهدف النوعي رقم ٢ طريقة المزاوجة بين صور الأشياء والكلمات التي تدل عليها . وفي هذا المستوى يجب أن يعطى لجميع التلاميذ تعليمات شفوية تقرر أنهم يجب عليهم أن يرسموا خطا يصل بين الكلمة والصورة التي تدل على معنى الكلمة . ومثل هذا السؤال يعطى قياسا مباشرا لقدرة الطفل على ربط كل كلمة بالصورة الملائمة . والعلاقة البسيطة بين الهدف وهذا الاختبار تدل على فائدة صياغة الاهداف في عبارات تدل على سلوك يمكن ملاحظته .

وبالنسبة لتقويم الهدف ٢ فان المعلم يطلب من تلاميذه الانتقال الى مواد جديدة وقراءتها قراءة صامتة . وبعد هذه القراءة توجه اليهم أسئلة يطلب منهم كتابة اجابات عليها وتتعلق هذه الاسئلة بمحتوى للفقرة . وهذا الاختبار قد يأخذ صورة أن يطلب من التلاميذ أن يكتبوا جملة في الاجابة على السؤال - وهو صورة مبسطة لاختبار المقال ، الا أن التلاميذ في هذا المستوى لا تتوقع لهم أن تكون عندهم سهولة كبيرة في الكتابة ، ولذلك فان بعض الاختبارات الموضوعية ينتج قياسا أكثر صدقا لفهم القراءة عندهم مثل أسئلة الاختبار من متعدد مثل .

في يوم العيد ذهبنا الى : الحقل

المدرسة

الحديقة

ومرة أخرى يجب أن تكون تعليمات هذه الأسئلة شفوية وتطلب من التلميذ أن يضع خطا أو يرسم دائرة حول الاجابة الصحيحة للسؤال .

وحيث أن القراءة هي مهارة اتصال ، فان جميع اهداف التعليم في هذا الموضوع تهتم اساسا بالفهم . وخاصة فهم معاني الكلمات . فمن المتوقع عن التلميذ أن يشرح معنى كلمة ، أو يختار مرادفا لها ، أو يربط بينها وبين

الصورة أو الشيء الذى تصفه • كما تربط الأهداف الأخرى بينهم الجمل أو الفقرات •

ومع تقدم الطفل فى صفوف المرحلة الابتدائية فإن التأكيد على التحصيل فى القراءة يتزايد اهتمامه بأنماط مختلفة وجوانب مختلفة من الفهم ، فتتطلب اختبارات الفهم من التلميذ أكثر من مجرد استدعاء أجزاء من معلومات الحقائق والتفاصيل •

تقويم التحصيل فى الحساب :

تتزايد أهمية المهارات الحسابية والفهم الحسابى فى الحضارة المعاصرة تزايداً مستمراً • ومن المهم لتلاميذ المرحلة الابتدائية أن يتلقوا تعليماً جيداً فى المهارات الأساسية فى الحساب وفى عناصر التفكير الكمية • وللمساعدة فى تحقيق هذا الهدف من المهم للمعلم أن يقوم بتقويم مستمر لتقدم كل تلميذ • وهذا مهم على وجه الخصوص فى الحساب حيث كل مهارة لاحقة تعتمد على المهارة السابقة ، وحيث يفشل التلميذ إذا فشل فى إتقان أى من المعلومات الأساسية مثل :

الهدف النوعى	الهدف المرحلى	الغرض العام
١ - أن يستطيع التلميذ أن يعد ويسجل العدد الصحيح لأشياء أو حروف توضحها الصور فى حدود ١٠ وحدات	١ - أن يستطيع التلميذ أن يعد حتى الرقم ١٠٠ •	أولاً : أن يستطيع الشخص المتعلم حل مشكلات العدد والحساب
٢ - أن يستطيع الطفل العد الى الرقم ١٠٠ بزيادة واحد •	ب - أن يعرف التلميذ حقائق الجمع الأساسية	
٢ - أن يستطيع التلميذ الاجابة على مسائل الجمع لأزواج من الأعداد التى تتكون من رقم واحد بحيث يكون مجموعها أقل من ١٠		

د - أن يستطيع التلميذ
فهم المصطلحات شرح معنى مصطلحات مثل
الحسابية الخاصة جمع وطرح ومجموع وأكثر
واقل .

هـ - أن يستطيع التلميذ أن
يختار اختبارا صحيحا للأشياء
التي تتحدد بمصطلحات مثل
الأول والثاني والثالث . الخ .

وجميع هذه الأهداف النوعية الخمسة يمكن تقويمها إما شفويا أو
تحريريا ، فالهدف الأول يمكن قياسه باستخدام أسئلة بسيطة يعطى فيها
للطفل أعدادا مختلفة من الأشياء عليه أن يعدها .

تعليمات (شفوية) : اكتب على الخط الموضوع أسفل كل صورة العدد
الذى يدل على عدد الحروف فى كل منها .

د د د	ح ح	ص ص	ص د
د د	ح ح	ص ص	د
د د د	ح ح	ص ص	
_____	_____	_____	_____

ويمكن تقويم الهدف الثانى بسؤال التلميذ أن يعد جهريا أو أن يكتب
الأعداد من ١ الى ١٠٠ .

والهدف الثالث يمكن قياسه بكفاءة أكثر بمجرد سؤال التلميذ أن يكتب
اجاباته على عدد من المسائل التى تتضمن الجمع البسيط :

٦	١	٥	٧	٢
٢ +	٨ +	٤ +	٢ +	٤ +
_____	_____	_____	_____	_____

والهدف الرابع يمكن تقويمه بسؤال التلميذ مباشرة أن يشرح هذه

المصطلحات شفويا أو تحريريا . ويوجد تعديل طفيف لهذا الهدف بحيث يتطلب من التلميذ أن يتعرف على صور أشياء أو عمليات يمكن أن توصف لفظيا كان يعطى له رسم لعدد من الكرات متفارقة الحجم ويطلب منه وضع خط تحت الكرة الأصغر .

والهدف الخامس يمكن تقويمه باستخدام أسئلة مماثلة ، كأن يعطى له رسم لعدد من الكرات المتساوية ويطلب منه وضع خط تحت الكرة الرابعة مثلا .

ان انواع الأسئلة التي ذكرناها حتى الآن تتعلق أساسا بقياس معلومات التلميذ وتذكره لحقائق خاصة أو مصطلحات نوعية . الا أن المعلم قد يهتم بأنماط أخرى من القدرات وخاصة تلك التي ذكرناها في تصنيف الأهداف التربوية (لبلوم) وتحت عناوين مثل الفهم والتطبيق .

مثال لترجمة الأعداد والمفاهيم احداها الى الآخر :

١ - ما الذي يعطى نفس الاجابة التي يعطيها ٥ x ٤ ؟

$$٤ + ٤ + ٤ + ٤ + ٤$$

٥

(د)

$$(١) \frac{٦}{٨} (ب) \frac{٥}{٨} (ج) \frac{٤}{٨} (د) \frac{٣}{٨}$$

ومن اهم الأهداف التعليمية التي تحتاج الى تقويمها في اختبارات الحساب هي القدرة على تطبيق معرفة المبادئ والأساليب أو الاجراءات (وتقاس باختبارات المسائل الحسابية من مختلف الأنواع) .

تقويم التحصيل في العلوم :

مع زيادة التأكيد على العلم في منهج المدرسة الابتدائية ككل ، نجد من

الطبيعى أن تسعى المدرسة الابتدائية الى زيادة مدى ما يتضمنه برنامجها من العلوم . والأهداف التى تسعى الى تحقيقها يجب أن تكون أكثر من مجرد معرفة حقائق معينة أو مصطلحات أو حفظ مبادئ وقوانين . وإنما يجب أن تتضمن قدرات مثل القدرة على اجراء تجارب بسيطة ، واشتقاق المبادئ والتميمات من البيانات أو اعطاء أمثلة للمبادئ المختلفة وتطبيق مثل هذه المبادئ فى شرح النتائج أو التنبؤ بها كما يجب أن تتعلق الأهداف باتجاهات وميول معينة .

وحيث أن أغلب معلمى المدرسة الابتدائية يعوزهم الاعداد السليم فى العلوم فإنهم يحتاجون الى المساعدة فى تحديد أنماط التعلم التى تنتج عن مقرر العلوم تحديدا واضحا وفى الحصول على أمثلة خاصة من القدرات المختلفة . ومن الأساليب المفيدة فى هذا الصدد اعداد وانتقاء اختبارات جيدة وأسئلة جيدة فى الاختبارات ، فكثير من القدرات يمكن التدليل عليها بطريقة جيدة عن طريق تطبيقها فى مواقف تعد أسئلة الاختبار أمثلة لها أو مواقف تتحول الى أسئلة اختبار . ومعرفة الطريقة التى يمكن بها تقويم هذه القدرات يساعد فى توضيح المعنى الحقيقى للقدرة عند العلم وتشجعه على تدريسها .

وفيما يلى مثال :

الغرض العام	الهدف المرحلى	الأهداف النوعية
أولا : على المواطن المتعلم أن يقدر مصادر الأمة .	(١) أن يفهم التلميذ أهمية المصادر الطبيعية مثل الماء والأرض والنبات والحيوانات	١ - أن يستطيع التلميذ أن يذكر أسباب ضرورة المحافظة على مصادرنا من الماء والأرض والنبات والحيوان .
ثانيا : يقيس المواطن المتعلم التقدم العلمى بإسهامه فى الرفاهية العامة .	(١) أن يشعر التلميذ ببعض الاسهامات الهامة التى يقدمها العلم فى مجال الرفاهية العامة	٢ - أن يستطيع التلميذ تحديد الطرق المختلفة للمحافظة على المصادر الطبيعية .
١ - أن يستطيع التلميذ أن يتعرف على اسهامات علماء مثل باستير واديسون وسولك ومركونى .		

ثالثا : تتوفر لدى (١) أن يفهم التلميذ ١ - أن يستطيع التلميذ المواطن المتعلم وسائل الطريقة العلمية في حل الدفاع ضد الدعاية • المشكلات • بالوصول الى النتائج من الأدلة والشواهد •

(ب) أن يصل التلميذ ٢ - عندما يطلب من الى قرارات على أساس التلميذ أن يتخذ قرارات في الأدلة والشواهد سواء مواقف فرضية في (مجالات في مسائل العلم أو العلوم وغيرها) فانه يصل المسائل غير المتعلقة الى تلك التي تدعمها الشواهد بالعلوم • والأدلة •

ان الهدفين النوعيين المشتقين من الغرض العام الأول يهتمان بمعرفة المبادئ والطرق • ويمكن تقويم الهدف الأول بمجرد سؤال التلميذ أن يذكر قائمة - اما شفويا أو كتابة - بعدد معين من الأسباب التي تتعلق بالمحافظة على كل مصدر من المصادر الطبيعية المذكورة •

ويمكن للهدف الثانى أن يصاغ على نحو يشبه الهدف الأول ، وبالتالي يمكن تقويمه بأن تطلب من التلميذ أن يذكر قائمة شفوية • ويمكن أيضا تقويم هذا الهدف باستخدام الاختبارات الموضوعية •

مثال : حتى يستطيع الفلاح أن يحصل من أرضه على محصول وفير عليه أن يزرع في حقله •

(١) نفس المحصول عاما بعد عام •

(ب) محصولا واحدا لثلاث سنوات متوالية ثم محصولا آخر للسنوات الثلاثة التالية :

(ج) محصولا واحدا لخمس سنوات ثم يليه محصول آخر للسنوات الخمس التالية •

ويمكن صياغة هدف رقم ١ بحيث يسمح للتلاميذ بانتقاء الأسباب

الهامة للمحافظة على المصادر الطبيعية ، وعلى هذا النحو يمكن تقويمه باستخدام أسئلة الاختيار من متعدد . أما الهدف النوعى المعتمد على الغرض العام الثانى فيمكن اختباره باستخدام أسئلة بسيطة مثل أسئلة المزاوجة بين انواع الاسهامات التى قام بها مختلف العلماء .

والأهداف المشتقة من الغرض العام الثالث أصعب فى تقويمها نوعا ما . ويمكن أن تفيد فى ذلك الاختبارات الشفوية واختبارات المقال ، حيث يعرض على التلميذ مواقف مشكلة معينة ثم يطلب منه اعطاء الاستنتاجات والحلول المناسبة . كما يمكن تقويم هذه الأهداف باستخدام أسئلة الاختيار من متعدد من النوع التفسيري كما هو مبين فى المثال الآتى :

مثال : قام أربعة تلاميذ بالاشتراك فى تجربة لاستنبات الفول كما يلى :
قام المعتز بمسح بعض حبوب الفول بقطعة من القطن ثم تركها جافة .
قامت مها بمسح بعض حبوب الفول بقطعة من القطن ثم تركتها مبللة .
قام خالد بوضع بعض حبوب الفول فى طين جاف وتركها جافة .
قام أحمد بوضع بعض حبوب الفول فى طين مبلل وتركها مبللة .

ومن نتائج هذه التجربة حاول تلاميذ الفصل الحكم على أى النتائج الآتية حول الصحيح :

- (أ) تنبت حبوب الفول اذا تركت فوق قطعة من القطن الجاف .
- (ب) تنبت الحبوب اذا تركت فوق قطعة من القطن المبلل .
- (ج) تنبت الحبوب اذا تركت فى الطين الجاف .
- (د) تنبت الحبوب اذا تركت فى الطين المبلل .
- (هـ) تحتاج الحبوب الى ان تظل مبللة حتى تنبت .
- (و) تحتاج الحبوب الى ان تظل جافة حتى تنبت .
- (ز) تحتاج الحبوب الى القرية والماء حتى تنبت .

وبالنسبة لكل من العبارات الآتية افترض انها تصف النتائج الحقيقية للتجربة وعليك ان تقرر أى الاستنتاجات السابقة هو الأقرب الى الصواب

إذا كانت النتيجة واحدة من هذه العبارات • حدد اختيارك بكتابة الحروف
أ ، ب ، ج ، د ، هـ ، و ، ز على الخط القصير أمام العبارات الخاصة بالنتيجة:

- ١ - حبوب المعتز ومها أنبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠
- ٢ - حبوب أحمد أنبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠
- ٣ - حبوب مها أنبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠
- ٤ - حبوب مها وأحمد أنبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠
- ٥ - حبوب خالد وأحمد أنبتت ولم تنبت حبوب الآخرين ٠٠٠٠
- ٦ - حبوب المعتز وخالد أنبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠

وبهذه الطريقة يمكن قياس بعض جوانب القدرة على التفكير العلمي في
طفل المرحلة الابتدائية •

الفصل الثانى عشر

الاختبارات المرجعة الى المحك

مقدمة :

يغلب على الاختبارات المنشورة أنها اختبارات معيار ، وأسباب غلبة هذا النوع من الاختبارات المنشورة المرجعة الى المعيار أنها مصممة ومعدة وموضوعة بواسطة آخرين غير مستعملها . وإنما مبنية بحيث تكون قيمتها الكبرى فى كيف يتباين أداء جماعات مختلفة عليها .

ولكن من الممكن ، مع هذا ، أن تفسر درجات بعض الاختبارات فى ضوء عدد من العناصر التى يجب عليها التلاميذ اجابة صحيحة بصرف النظر عن مقارنة هذا العدد بأداء الجماعة . وفى مثل هذه الحالة نقول ان الاختبار يمثل محكا أدائيا performance criterion ، ونقول ان التلاميذ الذين يجيبون على كل عنصر من عناصر الاختبار اجابة صحيحة ، او يجيبون اجابة صحيحة من عدد محدد سلفا من العناصر فى الاختبار هؤلاء التلاميذ يمكن ان نقول انهم استوفوا حق الأداء الكامل المطلوب من الاختبار . هذا الاتجاه هو الذى يسمى الترجيع الى المحك criterion-referenced أو الترجيع المحكى criterion-referencing .

—ويجب ملاحظة أننا لا نستطيع أن نسمى أى اختبار اختبارا مرجعا الى المحك لمجرد أنه ليس اختبارا مرجعا الى المعيار . ذلك لأن معنى الترجيع الى المحك أنه الأداء على الاختبار مربوط أو متصل بمراجع السلوك (*) behavior referents ، وأن يكون الاختبار مصمما ومنشأ على هذا الاساس (*) . أى أن الاختبار بحكم طريقة تصميمه ، يجب أن يمدنا بمعلومات

(*) النطاق : الاداءات التى يتطلبها الرمز ، واسى منها نختار ينود الاختبار .

عن قدرة التلميذ على القيام بأداءات معينة بصورة مطلقة in absolute terms
فإذا كان الاختبار اختبار جمع وطرح الكسور مثلا ، فيجب أن يكون لدينا
أساس ما لزعمنا أن التلميذ يعرف كيف يجمع ويطرح الكسور ، إذا استطاع
أن يجيب عن عناصر الاختبار كلها ، أما إذا لم يستطع أى تلميذ أن يتم
الاجابة ، وإذا كان للاختبار حقا محك أدائى سلوكى ، فانه من الممكن أن
نخلص الى انه ما من تلميذ استطاع أن يحقق محك الكفاءة proficiency
criterion للمهارة المقاسة .

مصادر المحك :

ولكن من أين يأتى المحك الذى يرجع اليه الاختبار ؟

يقترح كوكس وفارجاس (١٩٦٦) Cox & Vargos أن محكا رئيسيا
للترجيع هو أن التدريب يؤدي الى زيادة الكفاءة . فإذا أعطى اختبار لجماعة
ما قبل التدريب وبعده ، وكانت نتائج الجماعة بعد التدريب أفضل منها قبله
فان هذا يعنى أن الاختبار حساس لنواتج التدريب . أى انه من الممكن تفسير
نتائج الاختبار على أنها تدل على اكتساب المهارات والمعرفة التى قصد التدريب
الى انتاجها . مثل هذا الاختبار يكون اختبارا مرجعا الى المحك .

وانه من المؤسف حقا أن يكون عدد قليل جدا من الاختبارات التى
ينشئها المدرسون اختبارات مرجعة الى المحك يمكن أن تتصف بالقدرة والمتانة .
ذلك انه حتى يمكن اعتبار اختبار ما مرجعا الى المحك ، من الضرورى أن
نحاول أن نحدد العلاقة بين الأداء على هذا الاختبار وسلوك محكى
criterion behavior . وحتى يكون لدينا تبرير سليم لاعتبار اختبار ما
اختبارا مرجعا الى المحك ، فالتا لا بد أن نتبع الخطوات التالية فى بنائه
وترجييعه :

١ - يعد تخطيط محتوى content outline يذكر فيه المهارات
والمعرفة التى يعتبر الاختبار محاولة لقياسها .

٢ - يميز الأداءات ، أى المرامي القابلة للقياس measurable objectives التى سوف يكون المفحوص قادرا عليها . مع التسليم بأن هذا المفحوص قد اكتسب كفاءة فى المهارات والمعرفة المقاسة بالاختبار .

٢ - يميز المجال أو النطاق domain الذى يحدده كل مرمى ، ويكتب العناصر أو البنود فى الاختبار وفق التفاصيل المعينة فى ذلك المجال أو النطاق ، ثم يختار عشوائيا عنصرين لكل مرمى ليبقى منها الاختبار .

٤ - يتأكد من صحة حقيقة أن المهارات والمعرفة المقاسة بالاختبار هى فى الواقع لازمة prerequisite للمرامي الأدائية أو المرامي السلوكية المميزة المطلوبة فى الخطوة رقم ٢ . وقد يكون هذا هو أكثر جوانب هذه العملية تغيرا ، ذلك لأن عملية اثبات صدق الاختبار validation تبدأ بحكم الشخص الذاتى ، أى بإجراء صدق ظاهرى ، ثم تمتد لتتضمن حكم مجموعة من الخبراء أو تتضمن بيانات أو معطيات حقيقية حصلنا عليها من اعطاء الاختبار لمجموعة أثبتت أداؤها الكفاءة لتبقى ما اذا كانت لدى أفرادها المهارات والمعرفة الخاصة بالاختبار .

٥ - يقرر أو يحدد درجة قطع (أو درجة قاطعة) cutoff score أو درجة محك criterion score تعبر عن الأداء على الاختبار الذى يجب أن يحصل عليه المختبر ليظهر أنه قد أحرز كفاءة كافية فى المهارات والمعرفة ليكون قادرا على أداء ألوان السلوك المحسكية criterion behaviors

ملاحح الاختبارات مرجعية المحك

وعلى هذا فان أهم ملاحح الاختبارات المرجعة الى المحك هى :

١ - أنها مؤسسة على عدد من المرامي الأدائية أو السلوكية والتى تعتبر هذه الاختبارات محاولة لقياسها .

٢ - وأنها مصممة بحيث يكون لها درجة عالية من الملاءمة appropriateness لأنها قائمة على مرام . ومحتواها يرتبط بمراميها .

- ٢ - أنها تمثل عيقات من السلوك أو الأداء الواقعي الحقيقي .
٤ - أن الأداء عليها يمكن أن يفسر في ضوء درجات قاطعة محددة سلفا .

وأنه من الأهمية بمكان أن نوضح أنه ليست هناك فواصل قاطعة بين الاختبارات المرجعة إلى المعيار والاختبارات المرجعة إلى المحك . وفي الواقع من الممكن أن نفسر اختبارا ما متجهين أحد الاتجاهين أو كليهما على شرط أن تكون المرامي أو تصنيفات المحتوى المتميزة هي الأساس الذي استعمل في كتابة عناصر الاختبار . على أن الفرق الأساسي بين الاختبارات المرجعة إلى المحك والاختبارات المرجعة إلى المعيار هو أن الأولى مؤسسة على درجة قاطعة محددة سلفا، مفترضا أنها داخلة في تصميم الأداء نفسه intrinsic أو هي وثيقة الاتصال بذلك الأداء ، بينما النوع الأخير من الاختبارات مؤسس على أداء جماعة معيار ، أي أن تفسير درجة الاختبار أساسه خارجي extrinsic

انشاؤها واستعمالاتها

والترجيع إلى المحك أكثر مناسبة في تطبيقه في مجال اختبار المهارة ، وأشد صعوبة في تطبيقه لقياس السلوك المعقد مثل التفكير وحل المشكلة ، كما يذهب (Ebel, 1970a, Cronbach, 1977) وعلى الرغم من هذا فإن ثمة اهتماما وشغفا باستخدام الترجيع إلى المحك استخداما أوسع في قياس التحصيل المدرسي ، سواء بالنسبة إلى الاختبارات التي ينشئها المدرس ، أو الاختبارات المنشورة .

ويستخدم الترجيع إلى المحك عن طريق :

- ١ - وضع ، واختبار ، مجموعة من المرامي تمثل النواتج المرغوب فيها للتدريس .

٢ - تصميم عناصر أو بنود ، أو اختبارها بحيث نقيس كل مرمى قياسا دقيقا، أي تكون هذه العناصر أو البنود ممثلة لنطاق المرمى objective domain

٢ - التحديد المسبق presetting لمستويات أداء مقبولة .

٤ - اعطاء الاختبار لمجموعة من التلاميذ وتقويم أدائهم في ضوء عدد من المرامي التي يستطيعون أداء ما تتطلبه من أداء بنجاح .

ويتضمن الترجيع الى المحك مهمة كتابة أو وضع مرام وعناصر اختبار لقياس هذه المرامي . وبينما نجد هذا يسيرا على المدرسين في فصولهم ، فإن كتابة مرام وعناصر لتستخدم على نطاق أوسع ، لمنطقة أو محافظة مثلا ، أمر على درجة كبيرة من التعقيد ، وربما كانت مؤسسة اختبارية أكثر صلاحية وقدرة على القيام بها من المدرس . إلا أنه حتى المؤسسة الاختبارية ذاتها سوف تجد من الشاق أن تعد اختبارا مرجعا الى المحك على أساس قومي أو حتى اقليمي ، ذلك لأن مثل هذا الاتجاه سوف يضحى بتصويب الاختبار نحو حاجات نظام مدرسي بعينه . وإذا كانت الحقيقة هي أن المدخلات المحلية أكثر ضرورة للاستخدام المحلي ، وأن عددا جد قليل من الأفراد المحليين ، المدرسين في المدرسة مثلا ، لديهم الوقت ، أو المهارة لإنشاء اختبارات يمكن استخدامها في نطاق واسع ، هذه الحقيقة تعد كثيرا من الامكان والقيمة التطبيقية للاتجاه الى بناء اختبارات تحصيلية مرجعة الى المحك .

وعلى أية حال ، في الولايات المتحدة ، نجد في زيادة « بنوك » أو « مخازن » مرامي الأداء performance objectives ، وعناصر الاختبار وبنوده . وزيادة رغبة مديري المدارس في أن يكونوا قوائم مرامي الأداء لمناطقهم . ونمر شركات الاختبار في هذا المجال الخاص بالاختبارات المرجعة الى المحك ، كل هذا زاد من احتمال أن الاختبارات المرجعة الى المحك سوف تكون ميسرة للاستعمال في قياس التحصيل المدرسي .

ولهذه الاختبارات المحكية ، سيما يبنو ، ميزة السماح لكل منطقة تعليمية بأن توجه برنامجها التقريبي نحو أهدافها الخاصة ، وأن تعرف مدى إحراز هدفها بشكل مطلق وليس بشكل نسبي . فليس كافيا أن تقول إن محمدا تعلم أكثر من حامد ، ذلك لأنه مازال علينا أن نعرف عدد المرامي التي حققها كل منهما حتى تسمح بالتقدم نحو أهداف أخرى أكثر تركيبا . وسوف يقدم ناشرو الاختبارات مساعدة كبيرة للمدارس إذا أمدها بعناصر الاختبار التي

تتفق مع المرامى بحيث تكون كل مدرسة قادرة على ان تشكل اختبارها التحصيلى وتصوغه • موجهها نحو حاجات تلاميذها •

٥- تفسير درجات الاختبار المرجع الى المحك

درجة الاختبار ليست سوى رقم • ماذا تحمل هذه الدرجة من معنى ودلالة ؟ ماذا تكشف عنه ؟ هل ايدى التلميذ كفاءة فى المرامى المقاسة ؟ هل لديه قدرة ليتحرك الى مستوى أداء تال ؟ هل لديه القدرة على أداء مهارات على مستوى مطلوب لبداية العمل فى مهنة ما ؟

درجة الاختبار فى ذاتها لا تدلنا الا على القليل جدا ، وحتى تكون مفيدة يجب ان يعطى لها تفسير •

واذا كان الاختبار قائما على عدد من المرامى ، وكانت هذه المرامى ذاتها لها درجة من الصدق ، فان الدرجة على الاختبار سوف تخبرنا بشئ عن قدرة التلميذ على الأداء فى مجال الاختبار • ويجب ان يمدنا الاختبار المرجع الى المحك بمعلومات عن درجة كفاءة التلميذ فى المرامى التى يقيسها الاختبار • فالدرجة المرتفعة على الاختبار يجب ان تدل على احراز المرامى وتحصيلها ، بينما الدرجة المنخفضة تدل على نقص التحصيل • وقد تظن انه ربما كانت عناصر الاختبار سهلة جدا او ضعيفة جدا ، ومن ثم فانها تكون غير منصفة فى قياس التحصيل • ولكن اذا كان اختبارك قد استوفى محكات الملاءمة ، والصدق ، والوثبات ، فانه من الأرجح ان تكون عناصره مناسبة لقياس التحصيل ، ومن ثم يمكنك ان تحكم بان النجاح فى الاختبار يقيس احراز المرامى وتحصيلها •

وحتى يكون تفسير درجة الاختبار التحصيلى دقيقا وسليما ، ربما يكون من الأفضل ان تقيد نفسك بالدرجة الكلية على الاختبار • ولما كان الاختبار المرجع الى المحك أكثر شيها بمجموعة من الاختبارات المصغرة mini-tests يقيس كل منها مرمى واحدا • وكل منها يحتوى عنصرين ، وأحيانا أكثر من عنصرين ، لما كان هذا شأن الاختبار المرجع الى المحك ، فان التحصيل يمكن ان يوصف عن طريق ذكر المرامى الخاصة التى أظهر فيها كل تلميذ كفاءة

بحكم نجاحه في الاختبار المصغر المناسب أو المتصل بهذه المرامي . وعلى هذا ، فإنه لابد من أن تحتاج الى أن تقرر ما هي الدرجة المرتفعة التي يمكننا اعتبارها مرتفعة . يمكن أن تسجل لكل تلميذ اسم المرمى الذي إظهر فيه كفاءة ، ولا بأس - إذا كان هذا مرغوباً فيه - بإعطاء التلميذ ما يدل على مستوى الكفاءة الذي أحرزه فعلاً . فتفسيرية مثل هذا الاختبار قائمة أساساً لا على الدرجة الكلية بل على دلائل درجة الكفاءة التي أمكن إحرارها لكل مرمى من المرامي التي يقيسها الاختبار . ولما كانت المرامي تمثل اهتمامنا الحقيقي ، وكانت الاختبارات المرجعة الى المحك تقيس تحصيل هذه المرامي فانتنا نستطيع أن نركز نتائج اختبارنا وتفسيراتنا على ذلك التعلم الذي تريد أن تدفعه وترفعه .

الاختبارات التحصيلية المرجعة الى المحك

الترجييع الى المعيار norm-referencing واحد من أهم مكونات الاختبارات المقننة . والاختبارات التحصيلية المرجعة الى المعيار تفسر درجاتها بمقارنتها بدرجات حصل عليها بالفعل تلاميذ آخرون . أولئك الذين يكونون الجماعة المعيارية أو جماعة المعيار norming group وتقرر ما اذا كانت درجة ما منخفضة أو مرتفعة بمقارنتها بدرجات حصل عليها تلاميذ من مستوى الصف ذاته في الجماعة المعيارية .

ويمكن القول ان كل اختبار ينشئه المدرس هو بالامكان اختبار مرجع الى المحك ، لأنه مصمم لقياس درجة الكفاءة المنخفضة على عدد معين من المرامي . ومن العجيب ان كثيراً من المدرسين لا يبنون اختباراتهم حول مرام معينة ومحددة ، بل ان كثيرين منهم يصححون اختباراتهم كما لو كانت مرجعة الى المعيار بأن يقيموا تصحيحهم على أساس الأداء النسبي للمفحوص ، في حين أنه يجب تصحيح الاختبار في ضوء أداء على مرام أو عناصر ، أي على أساس الترجيع الى المحك .

الاهتمام الرئيسي في الاختبارات المرجعة الى المعيار هو : الى أي مدى يمكن أن يقارن فرد أو مجموعة أفراد مع آخرين ؟ والاجابات عن هذا السؤال

مفيدة جدا في ضمان تحقيق مستوى أدنى من الأداء النسبي في الفصل ، أو المدرسة ، أو المنطقة التعليمية .

أما الاختبارات المرجعة الى المحك فتسال : بأى شكل يسلك فرد أو مجموعة من الأفراد ؟ ما الذى يعرفه الفرد ذاته ؟ ما الذى تعرفه المجموعة ؟

ومن ثم فان الترجيع الى المعيار يعطى نتائج تعيينية ، summative أى أنها تخبرنا أين يقف الفرد أو مجموعة الأفراد ، بينما الترجيع الى المحك يعطينا نتائج تشكيلية formative ، أى أنها تخبرنا فى أى المجالات يمكن أن يوجه التدريس بحيث يتيسر تحصيل الكفاءة . الاختبارات المرجعة الى المحك تساعد المدرسة على تعرف monitor تقدم التلميذ . وتشخيص مواضع القوة والضعف ، ووصف التدريس وتوجيهه ، وخاصة العلاجى منه .

وبينما يستطيع المدرسون انشاء اختباراتهم التحصيلية المرجعة الى المحك ، فان الاختبارات المنشورة تعطيهـم درجة أرفع من ضبط نوع اداة الاختبار ، وامكانات افضل فى ملامح تقرير نتائج الاختبار ، وفى امكان المدرسين ان يفسروا هذه النتائج فى ضوء مراميهم التدريسية الخاصة .

ومن الممكن بناء اختبار تحصيلى مقنن ، معيارى ، يكون فى الوقت ذاته اختبارا مرجعا الى المحك . كيف ؟

مقارنة بين الاختبارات المرجعة الى المعيار ، والاختبارات المرجعة الى المحك

تعرف ان الاختبارات التحصيلية المقننة يتوافر لها ثلاث خصائص هي :

- ١ - ان عناصرها وينودها جربت ، وحلت ، وروجعت .
- و ٢ - انتشارها واتساع استعمالها واعادة استعمالها المقنن .
- و ٣ - وجود معايير لها واستعمال هذه المعايير فى تفسيرها .

وانه من الممكن بناء اختبار تحصيلى مقنن يكون فى الوقت ذاته مرجعا

الى المحك ، وذلك بتغيير واحدة من الخصائص الثلاث السابقة ، تلك هي الخاصية رقم (٢) ، بحيث تكون هكذا .

« وجود مرام للتفسير واستعمالها ، وتتضمن هذه المرامي محكات لتقويم الكفاءة ، فعندئذ نكون قد وضعنا الأساس . على الأقل من حيث التعريف ، لاختبار تحصيل « مقنن » ، يكون اختبارا تحصيليا مرجعا الى المحك »

هذا ، وان الفرق الوظيفي بين اختبار مرجع الى المحك (رك CR) وآخر مقنن او مرجع الى المعيار (رع NR) ذو ثلاث شعب :

الأولى : أن الاختبار المرجع الى المحك موجه توجيهيا خاصا الى مرام بينما الاختبار المرجع الى المعيار تكون درجاته أكثر كلية more global وشمولا ، ومن ثم فانك تحصل على درجات أكثر في الاختبار المرجع الى المحك ، (لأنه تحصل على درجة واحدة لكل مرمى) ، ولكن يوجد فيه عدد أقل من العناصر لكل درجة .

والثانية : درجات الاختبار المرجع الى المعيار يمكن أن تفسر عن طريق المعايير التي يعطيها ناشروه ، بينما درجات الاختبار المرجع الى المحك تتطلب منك أن تضع أنت نقطة أو درجة القطع التي تدل على الكفاءة المناسبة .

والثالثة : عناصر الاختبار المرجع الى المعيار تكتب بحيث تنتج أقصى قدر من التنوع في الأداء بين التلاميذ ، بينما الاختبار المرجع الى المحك تكتب عناصره بحيث تمثل نطاق المرامي كما يراه كاتب العنصر .

بعد أن نتعمق في هذه الفروق بين الاختبار المعياري NR (رع

والاختبار المحكي (رك) ، وبعد معرفة أن عناصر الاختبار المعياري قد كتبت مستعملة تخطيط محتوى content outline أو تصنيف مهارات skills classification ، فمن الممكن أن نفكر في استخدام اختبار

معيارى كما لو كان اختبار محك على شرط أن يذكر ناشره نتائج بمجال المحتوى أو مجال المهارة . وقد بدأ كبار ناشري الاختبارات منذ الآن يفعلون

هذا بالنسبة للاختبارات المعيارية أو المقتنة مثل اختبار أيوا Iowa Test
Metropolitan Achievement Tests of Basic Skills

وإن تفسيرات نتائج الاختبارات التحصيلية المرجعة إلى المحك يغلب
أن يكون مساعدا للمدرسين بصورة أكبر من تفسيرات الاختبارات المرجعة
إلى المعيار ، إلا أننا يجب أن ننتبه إلى نقطتين :

الأولى : من الضروري أن تكون المهارات والمحتويات المقاسة على وجه
الخصوص هي تلك التي تحتاج معلومات عنها ، وأنه يوجد عدد كاف من
العناصر لكل مهارة ، وذلك لتوفير الثبات .

والأخرى : لما كانت العناصر مكتوبة لأقصى تمييز بين الدرجات المرتفعة
والدرجات المنخفضة ، فإنك قد تحتاج إلى فحص درجات تلاميذك على
الاختبار لتحديد أي نقطة قطع ، أو درجة قطع ، تستخدمها للكفاءة المناسبة
بالنسبة إلى كل مهارة ، بدلا من استخدام نقطة قطع ثابتة مقرررة . وبمعنى
ما فإن هذا مازال صورة من صور الترجيع إلى المعيار ، ذلك لأن اختبار نقط
القطع سوف يقوم على الأداء النسبي relative performance للتلاميذ ، مثلا ،
عدد العناصر الجاية صوايا من ٧٠٪ من التلاميذ .

مراجع الباب الخامس

- ١ - أحمد زكى صالح : الأسس النفسية للتعليم الثانوى ، دار النهضة العربية ، ١٩٧٢ .
- ٢ - رايتسون ، ج . واين وآخرون (ترجمة محمد محمد عاشور وآخرين) : التقويم فى التربية الحديثة . مكتبة الأنجلو المصرية ١٩٦٥ .
- ٣ - رمزية الغريب : التقويم والقياس النفسى والتربوى . الأنجلو المصرية ، ١٩٧٠ .
- ٤ - سامى ابراهيم على : تقنين امتحانات البيانو . رسالة ماجستير ، المعهد العالى للتربية الموسيقية ، ١٩٧١ .
- ٥ - فان دالين ، ديوبولد (ترجمة محمد نبيل نوفل وآخرين) : مناهج البحث فى التربية وعلم النفس : مكتبة الأنجلو المصرية ، ١٩٦٩ .
- ٦ - فؤاد أبو حطب : دور التربية فى تنمية التفكير الابتكارى . مجلة الفكر المعاصر ، يونيو ١٩٧٠ .
- ٧ - ميكس ، روبرت ف . (ترجمة جابر عبد الحميد جابر وسعد عبد العزيز نادر) : الأهداف التربوية ، بغداد ، مطبعة العائى ، ١٩٦٧ .
- ٨ - نعيم عطية : التقييم التربوى الهادف ، بيروت ، منشورات دار الكتاب اللبنانى ، ١٩٧٠ .
9. Bloom. B. Taxonomy of Educational Objectives. London : Longmans, 1956.
10. De Cecco, J.P. The psychology of Learning and Instruction. Englewood Cliffs, N.J., 1968.
11. Mager, R.F. Preparing Objectives for Programmed Instruction. Palo Alto, Calif. : Varian Associates, 1961.

12. Ebel, B.L. Maximizing test validity in fixed time limits. *Educ. psychol. Measurements*, 1953, 8, 347-357.
13. Ebel, R.L. *Measuring Educational Achievement*, Prentice-Hall, 1965.
14. Gagne, R.M. *Conditions of Learning*, Holt., Rinehart, Winston, 1965.
15. Gronlund, N.E. *Measurement and Evaluation in Teaching*, Macmillan, 1965.
16. Linn, E.F. (ed.) *Educational Measurement*. Amer. Council on Educ., Washington, D.C. 1951.
17. Lindvall, C.M. *Testing and evaluation : An Introduction*. Harcourt, Brace, 1961.
18. Mager, R.F., Mc Cann, J. *Learner-Controlled Instruction* Palo Alto, Clif Varian Associates, 1961.
19. Thorndike, R.L. (ed.) *Educational Measurement* (2nd ed.) Amer. Council on Educ., D.C., 1971.

الباب السادس

مقاييس الشخصية

الفصل الثالث عشر

وسائل التقرير الذاتى

اشرنا فيما سبق الى تقسيم المقاييس النفسية بصفة عامة الى قسمين رئيسيين هما : مقاييس الأداء المميز ، ومقاييس اقصى الأداء ، وعرفنا ان هى النوع الاول من هذه المقاييس يكون الهدف من القياس هو معرفة ما يقوم به الشخص عادة ، بينما النوع الآخر من المقاييس يكون الهدف هو معرفة اقصى ما يمكن ان يقوم به الشخص فى ظروف استثارة ودافعية غير عادية .

ويمكن ان نضع تحت القسم الاول ، اى مقاييس الأداء المميز ، اختبارات الشخصية ، واختبارات الميول والاتجاهات بينما تدخل اختبارات القدرة باتواعها المختلفة تحت القسم الثانى وهو اختبارات اقصى الأداء . وفى هذا الفصل نتكلم عن قياس الشخصية باعتباره قياسا للسلوك أو الأداء المميز . وسوف نتبع . بصفة عامة ، التقسيم الذى حدده كرونباك (٢٧) لأنواع مقاييس الشخصية ، فندرجها تحت هذه الأقسام الثلاثة :

(١) قياس الشخصية عن طريق أحكام الآخرين والملاحظات المنظمة judgments and systematic observations ، وهى ما سيتناوله الفصل الرابع عشر .

(ب) قياس الشخصية عن طريق اختبارات الأداء performance tests وهى ما سيتناوله الفصل الخامس عشر .

(ج) قياس الشخصية عن طريق التقرير الذاتى self-report وهى ما سيتناوله الفصل الحالى .

وقياس الشخصية عن طريق التقرير الذاتى Self-report على الرغم من ان تعبير السلوك المميز الذى تكررت الاشارة اليه يحمل طابعا «سلوكيا» فان المذهب السلوكى لم يكن هو الذى اجهلهم فى نشأة وتطوير مقاييس الشخصية وانما يرجع الفضل فى هذا الى الاتجاه الظاهراتى

(الفينومونولوجى phenomenology) ، فهذا الاتجاه يتميز بأنه يركز على العالم، كما يدركه الشخص ، أو على العالم الذاتى للشخص ، بما فى ذلك ادراك الفرد لذاته ، وعلى أن السلوك انما يتحدد ، ولا يمكن أن يفهم الا فى ضوء هذا العالم الذاتى الداخلى ، كما يذهب هذا الاتجاه الى أن غالب مشكلات التوافق التى يتعرض لها الفرد ترجع الى ادراك الشخص للاحداث والأشخاص وليس لواقع هذه الاحداث أو الأشخاص فى ذاتهم . وبالتالى فإن العلاج النفسى ، وفق هذا الإتجاه يجب أن يركز على تصحيح ادراك الشخص .

نشير الى المدرسة الفينومونولوجية فى بداية كلامنا عن قياس الشخصية لأن أول مقاييس أو استفتاءات للشخصية كانت قد وضعت بهدف قياس هذا العالم الداخلى الذاتى من الادراك والشعور . فى العقد التاسع من القرن التاسع عشر وضع جالتون Francis Galton هذه المقاييس عندما احتاج الى طريقة يمكن أن يطبقها على عدد كبير من مفحوصيه أثناء دراساته للصور العقلية ، ثم استعمل هذا النوع من الاستفتاءات بعد ذلك ستانلى هول Stanely Hall فى دراساته الموسعة عن نمو المراهقين ولكن الاستفتاء كانت له وظيفة خاصة بالنسبة الى كل من جالتون وستانلى هول . فبالنسبة الى جالتون كان التقرير الذاتى يستعمل على أنه الوسيلة الوحيدة الممكنة للحصول على معلومات عن أمور وأحداث فى عقل المفحوص ، فى داخل رأسه ، بينما كان التقرير الذاتى بالنسبة الى ستانلى هول يستعمل بهدف تجنب الجهد الشاق والوقت الطويل الذى يمكن أن يستهلك فى الملاحظة المباشرة لسلوك المراهقين (٢٧ : ٤٦٤) .

اما أول استفتاء وضع بقصد قياس الشخص فقد كان قائمة وود ورث للبيانات الشخصية .
Woodworth Personal Data Sheet

وظهرت خلال الحرب العالمية الاولى عندما ظهرت الحاجة الى تمييز أولئك الجنود الذين عندهم قابلية للانهياء أثناء القتال ، ولم يكن يتوافر عدد كاف من المتخصصين لتحقيق هذا الهدف . لهذا وضع وود ورث قائمة من الأعراض التى يهتم الفاحص بالتعرف عليها لاستبعاد من تظهر عندهم ، وعندما أعطيت القائمة ميزت بالفعل أولئك الذين قرروا وجود عند كبير من

الأعراض عندهم ثم خضعوا لاختبار أكثر دقة ، وكانت لاختبار البيانات الشخصية هذه قيمة كبيرة في ذلك المجال .

ويعتبر اختبار روث الجد الأكبر لعدد كبير من استفتاءات التوافق ، وهي التي تتكون أساسا من قوائم من المشكلات أو الأعراض أو الشكاوى التي على المفحوص أن يختارها . إلا أن مثل هذه القوائم لا تدعى أنها تعطي وصفا دقيقا عن الشخصية وهي غالبا ما تعطي درجة واحدة تمثل مستوى التوافق . وهي تتكون من عناصر يمكن أن تميز بين أشخاص معروف أنهم غير متوافقين وبين آخرين معروف أنهم عابيون ، والاستعمال الرئيسي لمثل هذه المقاييس هو التعرف على أولئك الأشخاص الذين يحتاجون إلى رعاية وإرشاد نفسي ، وعلى هذا يمكن أن ننظر إليها على أنها أدوات فرز screening تساعد في تحديد الأشخاص الذين يختارون أعراضا وصفات وانتقادات تتصل بهم شخصيا ، وقد عدله بيرت وعربه واستعمله أحمد زكي صالح (١ : ٨٥١ - ٨٥٢) .

ولا يمكن اعتبار استفتاءات التوافق مقاييس دقيقة لسمات محددة ، فالمعلومات التي تمدنا بها معلومات ليس لها صفة العمق . لهذا فإنه في الفترة بين ١٩٢٠ - ١٩٤٥ . تغيرت النظرة إلى مثل هذه الاستفتاءات وذلك بتأثير المدرسة السلوكية واستنكارها للتأمل الباطني الذاتي الذي هو أساس اختبارات التوافق . فيجب أن يكون المقياس بديلا للملاحظة السلوك ، وعلى هذا يجب أن تؤكد أسئلته بشكل أكبر على ما يفعله الشخص لا على ما يشعر به أو يفكر فيه . وعلى هذا اتسعت مجالات الاختبارات بحيث تصف أكبر عدد ممكن من جوانب الشخصية . وتلخص الاستجابات بإعطاء درجات لعدد من السمات traits أو لأنماط الاستجابات ، وذلك لأنه كان ينظر إلى الشخصية في ذلك الوقت على أنها مجموعة أو حزمة من العادات ، وكان يحكم على الشخص حسب قوة سمات مثل : الثقة بالنفس ، الإصرار ، سهولة تكوين صداقات . الخ والسمات القوية هي السمات التي تصف استجابة يقوم بها الشخص عادة وبكثرة .

وقد كانت قائمة السمات أو أقسام السلوك التي تعطي لها درجات في الاختبارات تختار بشكل قطعي ودون خطة معينة في الاختبارات الأولى ،

قمتلا سمة مثل الثقة بالنفس كان مصدرها الخبرة اليرمية ، بينما قشتق سمة أخرى مثل الانطواء من نظرية من نظريات الشخصية . وقد ظهرت عشرات من الاختبارات بهذا الشكل ، تأخذ اللاحقة منها من السابقة عليها بضعة عبارات أو أسئلة وتضيف إليها وتصممها بحيث توضع فى تجمعات سمات مختلفة . الا أن أكثرها شيوعا هو اختبار بيرنرويتز للشخصية .
Bernreuter Personality Inventory.

وقد ترجمه الى العربية محمد عثمان نجأتى (*) وتعطى فيه الدرجات للدلالة على : النزعة العصابية ، الاكتفاء الذاتى ، الانطواء ، والسيطرة . ويتكون الاختبار من ٢٥ سؤالا يجيب عنها المفحوص بنعم أو لا . وقد وضعه بيرنرويتز على أساس أن سلوك الفرد فى موقف واحد قد يكشف عن سمات شخصية مختلفة ، وعلى هذا يمكن أن ينظر الى كل سؤال من أسئلة الاختبار على أنه يقيس أكثر من سمة من السمات ، وبالتالي يمكن أن تستخدم الاختبار الواحد للكشف عن أكثر من سمة تبعا لطريقة معالجة الاستجابات التى تحصل عليها (٥) .

وقد كانت دراسة فلاناجان (٢١) Flanagan اختبار بيرنرويتز بداية لاستعمال درجات السمات كما تحددها القواعد الاحصائية . لقد اعتقد فلاناجان أن السمات لا يمكن أن نطلق عليها أسماء خاصة بها الا اذا كانت الارتباطات فيما بينها منخفضة . وعلى هذا قام بحساب معاملات الارتباط لدرجات اختبار بيرنرويتز من تطبيقه على ٣٠٥ مراهق ، ووجد أن هذه السمات ليست مستقلة . فقد وجد أن الانطواء لا يختلف اقليلًا عن النزعة العصبية ، وعندما طبق التحليل العاملى وجد عاملين هما : الثقة بالنفس والميل الاجتماعى يمكن أن يفسر البيانات التى نحصل عليها من قياس السمات الأربعة الأصلية ، وانشأ مفاتيح خاصة لقياس العاملين أو السمتين اللتين وصل اليهما أى الثقة بالنفس والميل الاجتماعى . وقد ارتبطت درجات هاتين السمتين ارتباطا ضعيفا جدا ، وعلى هذا أصبح من الممكن أن نقول انهما يمثلان جانبين مستقلين من التقرير الذاتى .

(*) لا تتوافر لدينا بيانات سيكومترية عن هذا الاستخبار فى البيئة المصرية .

بعد هذا حدث تطور آخر له مغزاه ، ذلك هو اخضاع نظريات الشخصية تماما لبحث احصائي عن ابعاد dimensions يمكن ان تلخص الشخصية .
فقد أدت الارتباطات الداخلية بين عناصر الاختبارات الى ان اقترح جيلفورد مثلا ، ان الانطواء يمكن ان نقسمه الى انطواء اجتماعي (ج) ، وانطواء فكري (ف) ، واكتئاب (ك) ، والتقلب المزاجي (ت) ، وضبط للنفس (ض) . وعلى هذا انشا الاختبار الذي اسماه : اختبار العوامل ج - ف - ك - ت - ض Inventory of Factors S - 'I' - D - C - R.

ثم اُضيف بعد هذا ثمانية جوانب أخرى للشخصية . ولم تكن مقاييس جيلفورد مستقلة ، ولهذا أعاد بعض الباحثين تنظيم هذه المقاييس في أنماط تجعلها أكثر كفاءة . من هؤلاء مثلا ترستون الذي استطاع ان يفسر كثيرا من المعلومات في مقاييس جيلفورد الثلاثة عشرة بسبعة عوامل فقط وأعاد تسميتها بحيث أصبحت هكذا التأمل ، الميل الاجتماعي ، الاتزان الانفعالي ، الحيوية أو الذكورة ، السيطرة ، النشاط ، الاندفاع ، وقد اقتبس منه أحمد زكي صالح اختبار الصفات الانفعالية .

وهكذا نجد أن واحدا من علماء النفس يصنف عناصر الاختبار التي تقيس السمات بشكل دقيق تفصيلي ، بينما يأتي التالي ليضعها في مجموعات صغيرة . ثم يأتي آخر فيعيد تقسيم المجموعات بطريقة جديدة ، ويعطي كلا منها أسماء جديدة للعوامل التي يصل اليها . وبهذا لا تنتهي هذه اللعبة ، كما يعبر كرونباك (٢٧ : ٤٦٧) عن رأيه أنه ما لم ترتبط قوائم السمات بنظرية محددة وبمحكات خارجية ، فإن عملية الاختبار لا يمكن إلا أن تكون عملية تفضيلية شخصية . وليس هناك اتفاق في الوقت الحاضر بين أصحاب التحليل العامل ، بالنسبة الى عدد العوامل التي تم تحديدها بشكل دقيق ، أو افضل تنظيم لها ، أو انصب الأسماء لها .

أما الاتجاه الأحدث في قياس الشخصية بصفة عامة فهو تحديد المفاهيم والتكوينات النظرية Constructs على أساس نظرية للشخصية ثم اعداد العناصر في الاختبارات بصورة تؤدي الى الحصول على معلومات عن هذه المفاهيم والتكوينات . وقد كان لنظرية يوتج Jung عن الشخصية تأثير في الدراسات الأولى عن الانطواء . ولكن هذه النظرية لم تتعد اقتراح بعض

عناصر اختبارات وضعت لجرد المحاولة . أما الآن فهناك دراسات واسعة عن اختبار مايرز - برجز Myers-Briggs Inventory والتي اشتقت عناصره ومفاتيح تصحيحه صراحة من نظرية يونج (٥١) . وتوجد مقاييس أخرى تمثل هذا الاتجاه منها مقياس التفضيل الشخصي لادوارد Edwards Personal Preference Schedule .

والذي يشتق من نظرية موري Murray عن الحاجات الخمس عشرة . ومن أمثلة هذه الحاجات : الانجاز ، والخضوع ، والنظام ، والعرض ، والاستقلال الذاتي ، السيطرة ، التأمل النفسي (المقصود بهذه الحاجة تحليل دوافع ومشاعر الفرد الذاتية وعند الآخرين) ، ويتكون المقياس من ٢١٠ زوجا من العبارات ، وعلى المقصود أن يختار عبارة من كل زوج على أنها أكثر تمثيلا لخصائص شخصيته ، ومن أمثلة هذه الأزواج .

(١) أحب أن أتحدث عن نفسي مع الآخرين .

(ب) أحب أن أعمل لهدف معين حددته بنفسي .

(١) أشعر بالاكتمال عندما أفشل في أمر ما .

(ب) أشعر بالاضطراب عندما أتحدث إلى جماعة .

ويهدف هذا المقياس إلى تقدير القوة النسبية للخمسة عشر متغيرا المرتبطة بالحاجات .

ويمكن تقويم الخمسة عشر درجة المرتبطة بالحاجات على أساس معايير مئوية وثائية لطلاب وطالبات الجامعة ، وهذه المعايير تقوم على عدد ٧٤٩ طالبا و ٧٦٠ طالبة موزعين على جميع أنحاء الولايات المتحدة ، وقد اضيفت معايير مئوية للكبار ينتمون إلى بيئات حضرية وريفية .

أما عن ثبات المقياس فقد ذكر دليل الاختبار معاملات ثبات لمقاييسه الفرعية الخمس عشرة معتمدة على إعادة الاختبار بعد أسبوع واحد تتراوح بين ٠.٧٤ ، ٠.٨٨ ، بينما معاملات الثبات المعتمدة على التجزئة النصفية تتراوح بين ٠.٦٠ ، ٠.٨٧ .

أما بالنسبة الى صدق مقياس ادواردز ، فعلى الرغم من قلة البيانات الواردة عنها فى دليل الاختبار ذاته ، فإن أبحاثا كثيرة مستقلة حول صدقه قد نشرت . الا أنها أعطت نتائج متعارضة ، وتذهب انستازى (١٨ : ٤٥٤) الى أن مقياس التفضيل الشخصى له مزايا كثيرة الا أنه محتاج الى أمرين :

أولا - مراجعة للتخلص من بعض جوانب الضعف الفنية فيه ، وبصفة خاصة فيما يتصل بشكل العناصر وتفسير الدرجة ، وبالثانى ، دراسات صدق تجرى بدقة مستفيدة من أساليب تحليل أنماط الدرجات مناسبة لدرجات هذا المقياس التى لا تتحدد فيها قوة الحاجة بصورة مطلقة ولكن بالنسبة الى قوة الحاجات الأخرى عند الفرد ، وعلى هذا فإن إطار المرجع فى هذا النوع من الدرجات هو الفرد ذاته وليس عينة التقنين : وقد يكون شخصان متفان فى درجاتهم على المقياس الا أنهما يختلفان اختلافا جوهريا فى القوة المطلقة لحاجاتهم .

وقد عرب هذا المقياس جابر عبد الحميد (١٧ : ٤٧٨ - ٤٨٠) وحسب ثبات الاختبار بطريقة التجزئة النصفية، وذلك بتطبيقه على طلاب الجامعة، وكان يتراوح بين ٠.٢٤ ، ٠.٧٧ وقد استعمل هذا المقياس فى بعض الدراسات المقارنة بين الثقافات وغيرها .

ومن المقاييس الأخرى التى اعتمدت على نظرية مقياس تيلور للقلق الظاهر Tayler Manifest Anxiety Scale وقد ارتبط فى تصميمه بالبحث فى نظرية هل - سبنس Hull-Spence السلوكية ، وسوف نعود اليه بعد دراستنا لاختبار مينسوتا المتعدد الأوجه الذى توليه اهتماما خاصا .

اختبار مينسوتا المتعدد الأوجه للشخصية :

Minnesota Multiphasic Personality Inventory (MMPI)

يعتبر هذا الاختبار من أهم الأدوات التى وضعت لمقياس الشخصية عن طريق التقرير الذاتى ، لهذا سوف نتناوله بشيء من التفصيل . نجد لهذا (التقييم النفسى)

الاختبار مكانا بين اختبارات الشخصية تشبه مكان اختبار سترونج Strong بين اختبارات الميول ، ذلك لأنه يشارك اختبار سترونج فى الأساس التجريبي الذى صمم وفقا له ، وكذلك يشبهه من حيث تعرضه لبحث عميق وطويل من جانب واضعيه . وقد ظهر هذا الاختبار فى وقت كانت الحاجة اليه شديدة ، وكان الاعتماد عليه كبيرا اثناء وبعد الحرب العالمية الثانية التى شهدت توسعا كبيرا فى علم النفس المرضى والعلاج النفسى . أى ان نشأته كانت أصلا نشأة اكلينيكية وهذا مما ساعد على انتشاره ، وكذلك ساعد على انتشاره أيضا بعض الأمور الفنية المبتكرة فيه ومنذ نشره فى عام ١٩٤٠ ونشر دليله فى عام ١٩٤٣ فإنه استثار فيضا من البحوث التى استخدمته .

وكان صميم الاختبار فى الأصل على يد عالم نفسى هو ستارك هاثاوى Starke Hathaway وطبيب نفسى هو ج . س . ماكنلى J.C. Mckinley (٢٥) وكان الهدف من هذا التصميم هو المساعدة فى تشخيص المرضى النفسيين وتحليل السمات التى يميز الشذوذ النفسى الذى يؤدى الى العجز .

يتكون الاختبار من ٥٥٠ عبارة يسأل المفحوص أن يصنفها وفق أحد اقسام ثلاثة ، أى يجيب عنها اما بنعم او لا ، ولا يستطيع ان اجزم او لا ادرى وفى الصورة الفردية للاختبار كانت العبارات تطبع على بطاقات منفصلة وعلى المفحوص ان يصنفها فى ثلاث مجموعات ، وبعد ذلك طبعت العبارات فى كتيب وتسجل الاستجابات فى ورقة الاجابة . وقد صممت الصورتان الفردية والجماعية للكبار ، أى من ١٦ سنة وما بعدها . الا انهما استخدما بنجاح مع مرافقين اصغر من هذا السن . وبصفة عامة يفضل استخدام الصورة الفردية وخاصة عند الأشخاص المضطربين أو من ذوى المستوى التعليمى أو العلى المنخفض .

وتتنوع عناصر الاختبار المتعدد الأوجه للشخصية وتوسع من حيث محتواها فتشمل مجالات مثل . الصحة ، والأعراض الجسمية - النفسية ، والاضطرابات العصبية ، والاتجاهات نحو الجنس ، والسياسة ، والدين ، والمسائل التربوية والمهنية والعائلية ، كما تشمل كثيرا من المظاهر والأعراض العصابية والذهانية المعروفة مثل المخاوف المرضية ، والاضطهاد ، ونزعات تعذيب الذات وتعذيب الآخرين وغيرها .

وعندما نشر الاختبار فى صورته الأولى كان يعطى درجات على تسعة مقاييس كلينيكية ، ويتكون كل مقياس منها من عناصر ميزت بين جماعة مرضية محدودة وبين جماعة ضابطة عادية ، وعلى هذا فان هذه المقاييس قد انشئت بطريقة تجريبية ، وذلك عن طريق تقسيم العناصر وفق مداه معين وهو التشخيص النفسى المعروف ووفق هذه الطريقة اعدت المقاييس التالية :

1. HS : Hypochondriasis	يسمى	١ - توهم المرض - ه س
2. D : Depression	ثلاثى	٢ - الاكتئاب - د
3. Hy : Hysteria	العصاب	٣ - الهستيريا - ه ي
4. Pd : Psychopathic deviation	ب د	٤ - الانحراف السيکوباتى - ب د
5. Mf : Masculinity-femininity	م ف	٥ - الذكورة - الأنوثة - م ف
6. Pa : Paranoia		٦ - البارانويا - ب ا
7. Pt : Psychasthenia		٧ - السيکائينيا - ب ت
8. Sc : Schizophrenia		٨ - الفصام - س ك
9. Ma : Hypomania		٩ - الهوس الخفيف - م ا

وفىما يلى تعريف مختصر بما يدل عليه كل من هذه المقاييس (١٥) :

١٥٢ - ١٦٥) .

١ - توهم المرض : وهو مقياس للاهتمام الزائد بالوظائف الجسمية ، والقلق الذى لا مبرر على الصحة .

٢ - الاكتئاب : ويرتبط هذا المقياس بالشعور بالانقباض والتشاؤم والاحساس بعدم الأهمية الذاتية .

٣ - الهستيريا : ويقاس هذا المقياس درجة تشابه المفحوص بالمرضى الذين تظهر عندهم أعراض تتخذ صورة شكاوى عامة أو محددة مثل العمى أو الشلل أو القىء أو الاضطرابات المعوية . أو الأعراض القلبية .

٤ - الانحراف السيکوباتى : وهو مقياس لدرجة تشابه المفحوص بجماعة السيکوباتيين الذين تتبثل الصعوبة الرئيسية عندهم ، فى نقص

الاستجابة الانفعالية العميقة وقلة الاهتمام بمشاعر الآخرين ، وعدم المبالاة بالمعايير الاجتماعية .

٥ - الذكورة - الأنوثة : وهو مقياس للذكورة والأنوثة فى أنماط الاهتمامات . وقد استخرجت عبارات هذا المقياس نتيجة مقارنة استجابات الذكورة من ذوى الاهتمامات الذكرية باستجابات الذكور من ذوى الاهتمامات الأنثوية ، أما الفروق بين استجابات الذكور والإناث فقد درسها واضعها الاختبار إلا أن هذه الفروق لعبت دورا صغيرا نسبيا فى هذا المقياس . وتدل الدرجات المرتفعة على هذا المقياس على غلبة أنماط الاهتمامات الخاصة بالجنس الآخر .

٦ - البارنوايا : وهذا مقياس للتشكك الزائد ، والحساسية المفرطة ، وهواجس الاضطهاد .

٧ - السيكاثينيا : وهو مقياس للمخاوف المرضية والسلوك القهرى ، الذى قد يكون صريح مثل التكرار الزائد لغسل اليدين ، أو ضمنا كما يتمثل فى عدم القدرة على التحرر أو الهروب من فكرة متسلطة .

٨ - الفصام : ويكشف هذا المقياس عن التفكير أو السلوك الخلطى الشاذ الذى لا يتفق مع الاتصال بالواقع وفهمه .

٩ - الهوس الخفيف : وهو مقياس للنشاط الزائدة فى الفكر والعمل ، وعدم التركيز الطويل ، والانشغال بعدد كبير من الأمور والمشروعات فى وقت واحد .

وتعطى بيانات جماعة التقنين من الإحاديين درجة معيارية (ت) يمكن تحويل للدرجات الخام إليها بحيث تخطط النتائج على بطاقة صحيفة نفسية profile sheet وتتركز الأهمية الأولى على الدرجات التى تزيد عن ٧٠ والدرجة ٥٠ هى متوسط جماعة المرجع) . وهذه الدرجة الفاصلة ، أى ٧٠ ، هى درجة مختارة على أساس قطعى اعتسافى ، أى ليس لها أساس أو تبرير نظرى . وعلى أى حال نجد أن مفسرى نتائج الاختبار يعيلون الى

الاهتمام بكل الارتفاعات ، أو القمم ، في بطاقة الصحة النفسية ، سواء كانت هذه الارتفاعات تتخطى خط الـ ٧٠ درجة أو لا تتخطاه .

مقاييس الضبط : ومن العلامات المميزة والخاصة بمقاييس ميخورتا المتعدد الأوجه استعماله لأربعة مقاييس تسمى مقاييس الصدق ، والصدق هنا معناه الضبط Control لأنها في الواقع تمثل ضوابط للاعمال ، وسوء القهم ، والتحليل ، ولتأثير انماط الاستجابات والاتجاهات نحو أخذ الاختبارات ، وهذه المقاييس هي :

١ - الأول ؟ ، وتمثل درجة هذا المقياس مجموع العناصر التي أجاب عنها المفحوص بالا أستطيع أن أجزم ، ومن المرغوب فيه أن تكون هذه الدرجة أقل ما يمكن .

٢ - الثاني ل L ، ويمثل الكذب ، ويعتمد على مجموعة من العناصر التي تجعل المفحوص يبدو في شكل مقبول اجتماعيا ، ولكن ليس من المحتمل أن تكون الاجابة عنها صادقة في الاتجاه المرغوب فيه . ومن امثلة هذا العنصر :

لا أقول الصدق دائما .

وعلى الرغم من أن الاجابة الصحيحة المعتادة تكون نعم ، الا أن الاجابة المقبولة اجتماعيا تكون لا ، ومن المفترض أن الشخص الذي يريد أن يظهر نفسه في صورة مقبولة يحصل على درجة مرتفعة في المقياس ل عن طريق تحريف استجاباته لعبارات المقياس ، وتبدل الدرجات العالية على هذا المقياس على أن الاجابات لا يمكن الثقة بها ، ولكنها لا تدل على أن المفحوص يعتمد الكذب .

٣ - الثالث ف F ، وهو مقياس الخطأ ، ويتكون من العناصر

التي تدر أن أجيب من مجموعة التقنين في نفس الاتجاه الذي أجابه فيها المفحوص ، وعلى الرغم من أن هذه العناصر لا يمكن أن تمثل نمطا معينا من الشذوذ النفسي فإنها تعبر عن سلوك غير مرغوب فيه ، وعلى هذا فإنه من غير المحتمل في الواقع أن يعبر مفحوص واحد عن كل أو معظم هذه الأعراض

وترتفع الدرجة على هذا المقياس عندما لا يستطيع الفحوص أن يعطى اجابة مميزة لسبب من الأسباب كان لا يستطيع القراءة او الفهم بدرجة مناسبة ، او الاهمال ، سواء عن قصد أم غير قصد ، كما قد يرجع ارتفاع الدرجة الى الخطأ فى التصحيح ، وبصفة عامة كلما زادت الدرجة على هذا المقياس زاد احتمال أن بعض العوامل تدخلت لتقلل من صدق الصحيفة النفسية (أو البروفيل) .

٤ - الرابع ك K ويمثل مقياس التصحيح ، وفى هذا المقياس استعمال لتركيبية أخرى من العناصر المختارة ، وتعطى درجة هذا المقياس دلالة على الاتجاه نحو أخذ الاختبار ، وهى متصلة بكلا المقياسين السابقين ل ، ف . ولكن يعتقد أنها أكثر خفاء منهما وتدل الدرجة العالية على هذا المقياس على اتجاه دفاعى defensiveness او على محاولة المخادعة فى الاتجاه الحسن ، « Fake good » او التحريف المقصود نحو الطرف السوى . كما تدل الدرجة المنخفضة على الصراحة المفرطة ، ونقد الذات ، او المحاولة المتعمدة للمخادعة فى الاتجاه السيئ fake bad او الميل الى اظهار النفس بالمظهر غير السوى .

ومن المهم أن نشير الى أن استعمال مقاييس الضبط الأربعة هذه ليس مقننا تماما ولكنه يترك جزئيا لحكم مستعمل الاختبار ، وكذلك نشير الى أن هذه المقاييس تخضع باستمرار لاعادة التحديد والمراجعة فى ضوء البحوث الجديدة .

كانت تلك هى المقاييس الاكلينيكية التسعة الأساسية ثم المقاييس الأربعة الأخرى الخاصة بضبط الاختبار ، الا أنه منذ نشر الاختبار أنشئ عدد كبير جدا من المقاييس الجديدة بواسطة باحثين آخرين لم يشاركوا فى وضع المقاييس الأصلية ، من هذه المقاييس الجديدة مقياس عن الانطواء الاجتماعى (س ي) ، Social Introversion (Si) وقد أضيف الى المقاييس التسعة الأصلية وهو مضمن فى الاختبار ورقمه (صفر) . ويقاس هذا المقياس النزعة الى الانزواء والابتعاد عن الاتصال الاجتماعى بالآخرين .

وبهذا فإن الاختبار فى صورته الحالية يعطى ١٤ درجة على الأربعة عشر مقياس ، منها التسعة الأصلية ، ومقياس الانطواء الاجتماعى ، ثم مقياس الضبط الأربعة .

ومعايير المجموعة الضابطة الأصلية والتي تتكون من حوالى ٧٠٠ فردا مذكورة فى صورة درجات ثنائية (ت) ٠ أو درجات معيارية بمتوسط ٥٠ وانحراف معيارى قدره ١٠ . وتستعمل هذه الدرجات المعيارية فى رسم الصحيفة النفسية ، و أى درجة تعادل ٧٠ أو تزيد عنها ، أى تقع فوق المتوسط بدرجتين انحراف معيارى ، ينظر اليها بصفة عامة على انها تميز انحرافا مرضيا ومما يجب ملاحظته ان المعنى الاكلينيكي لنفس الدرجة قد يختلف من مقياس الى مقياس آخر . فمثلا الدرجة ٧٥ على مقياس توهم المرض وعلى مقياس الفصام لا تدل على نفس الدرجة من حدة الانحراف .

اما تفسير الدرجات المنخفضة على الصحيفة النفسية للاختبار فليس على درجة كافية من الوضوح . اذ يعتقد الاكلينيكيون أن وقوع درجات أقل من ٥٠ كثيرا قد يكون له مغزى تشخيصى . ولكن لم توضع بعد تفسيرات منظمة لهذا .

على أى حال ، فإن هناك قدرا من نتائج البحوث التى أجريت على هذا الاختبار يوضح انه بصفة عامة كلما زاد عدد وحجم الدرجات المنحرفة على الاختبار ، زاد احتمال أن يكون الشخص مصابا باضطراب شديد .

ومن الواضح أن الاستعمال الرئيسى للاختبار المتعدد الأوجه هو التشخيص التمايز أو الفارق differential diagnosis عندما يستعمل فى هذا المجال لا بد أن يمارس مستعمله منتهى الحذر ولا يلجأ الى التفسير الحرفى للمقاييس الاكلينيكية فيه ، لأن الأمر أكثر تعقيدا من تلك الأسماء المعطاة للمقاييس الأصلية . فليس معنى الدرجة المرتفعة على مقياس الفصام مثلا أن الشخص مصاب بالفصام . فقد يحصل شخص عادى على نفس الدرجة . بينما قد يحصل المصابون بالفصام فعلا على درجات مرتفعة على مقاييس أخرى . ولكى يستبعد احتمال سوء التفسير للدرجات على المقياس الواحد فإن أرقاما من صفر الى ٩ حلت محل أسماء المقاييس فى طبعات

اخيرة من الاختبار . ولكى يسهل النظر فى نمط الدرجات او الصحيفة النفسية وفحصها وضع نظام ترميز عددي للصحيفة النفسية . وفى هذا النظام يدل تتابع وترتيب ارقام المقاييس ، بنظرة سريعة ، على النقط المرتفعة والمنخفضة فى الصحيفة النفسية لشخص ما . فمثلا الرمز ٩٤ - ٢ يدل على درجة مرتفعة على المقياس رقم ٤ (ب د) وعلى المقياس رقم ٩ (م ا) بترتيب متناقص ، أى أن الدرجة على المقياس الأول أعلى منها على المقياس الثانى ، كما يدل هذا الرمز على درجة منخفضة على المقياس رقم ٢ (د) .

والواقع أن أسماء المقاييس التى استبعدت كانت تستند الى تقسيم اكلينيكي يقوم على أساس نظري مشكوك فيه على الرغم من شيوعه وقت وضع الاختبار ، فالتقسيمات المستعملة والتى وضعت لها مقاييس الاختبار كانت محل شك ومراجعة فى علم النفس المرضى ، أما فى المجال العلمى الواقعى فان هذه التقسيمات لم تعد مقبولة .

ومن دواعى الحذر فى تفسير درجات المقاييس ان بعض هذه المقاييس لا تتمتع بثبات مرتفع . وتعتبر هذه نقطة ضعف فى الاختبار . لأن فاعلية أى تحليل للصحيفة النفسية نقل بسبب اخطاء الصدفة فى الدرجات التى تعتمد عليها فى هذا التحليل . فاذا كانت درجات المقاييس المفردة غير ثابتة ومتداخلة أى تترابط فيما بينها ترابطا عاليا فان كثيرا من الاختلافات فيما بين الدرجات ربما تكون ناتجة عن الصدفة .

وتتراوح معاملات ثبات الاختبار الناتجة عن اعادة الاختبار ، بالنسبة الى الراشدين الأسوياء وغير الأسوياء بين ٥٠ و ٩٠ . وكانت الفترة بين الاختبار واعادته تتراوح بين بضعة أيام الى عام . الا أن هناك دراسة أجريت على طلاب الجامعة كانت معاملات الثبات فيها عموما أقل من السابقة ، كما اظهرت الدراسة نفسها تنوعا كبيرا بين مقاييس الاختبار كانت تتراوح بين ٥٠ لمقياس ب ١ ، ٨١ - لمقياس ب ٢ .

وثمة عيب آخر فى الاختبار يتصل بمدى تمثيل عينة التقنين . فالدرجات المعيارية التى اشتقت منها تنظيمات الصحائف النفسية موزوعة فى ضوء جماعة ضابطة تتكون من ٧٠٠ راشد هم الذين اختبروا فى التقنين الأسمى

للاختبار . وتبدو هذه العينة قاصرة جدا اذا قررنت بعينات التقنين الأخرى الأكثر شمولا والتي استعملت ، مثلا في تقنين اختبارات القدرة وقد أكدت أبحاث كثيرة ضرورة أن تكون المعايير متنوعة تنوعا كبيرا في جماعات أصل مختلفة ، ومن هذه الأبحاث ما وضع أن متوسطات درجات طلبة الكليات كانت فوق ٥٠ بشكل مطرد على بعض المقاييس ، كما بينت دراسة أخرى أن ٢٩٪ من طلبة الكليات حصلوا على درجات فوق ٧٠ على مقياس أو أكثر من مقاييس الاختبار .

ويتأثر الاختبار المتعدد الأوجه ، مثل سائر اختبارات الشخصية ، بالاختلافات الثقافية العامة بين المجتمعات بل وبالمستويات الثقافية الفرعية داخل المجتمع . وقد أظهرت دراسات كثيرة أجريت في بلاد مختلفة ارتفاعا ملحوظا في بعض مقاييس الاختبار عندما وضعت الصحائف النفسية على المعايير الأمريكية الأصلية . وأن تفسير مثل هذه الاختلافات في ثقافة المجتمع أو ثقافته الفرعية يتطلب معرفة بخصائص هذه الثقافات وظروفها ، فقد يكون الاختلاف في درجات الاختبار المتعدد الأوجه راجعا الى اختلاف الأفراد في تفسير عبارات الاختبار أو تعليمات الاختبار ، فمثلا قد يرجع الارتفاع في بعض المقاييس عند بعض الجماعات الى اعتقادات متوارثة نحو التواضع وتقليل شأن الذات . وكذلك يتأثر الدرجة بالاختلافات الثقافية يتصل بالسلوك الذي يعتبر مقبولا اجتماعيا . إلا أن الدرجات المرتفعة على مقاييس هذا الاختبار ، بالنسبة الى بعض المجتمعات والجماعات ، قد تدل على وجود مشكلات انفعالية حقيقية يمكن إرجاعها الى أساليب تنشئة الطفل ، أو الصراعات في الأدوار الاجتماعية ، أو غير ذلك من أساليب الاضطراب الانفعالي .

أما عن الترجمة العربية لاختبار الشخصية المتعدد الأوجه فقد أعدها عطية هنا ، ومحمد عماد الدين اسماعيل ، ولويس كامل مليكه . ووضعت له معايير من البيئة المحلية ، وحسبت معاملات الثبات للمقاييس بطريقة للتجزئة النصفية للمقياس ، وبطريقة كيودر - ريتشاردسون . وقد استخدم الاختبار في صورته العربية في عدد من الدراسات (١٠ : ٢١٦ - ٢٢٤) .

اختبارات مشتقة من اختبار الشخصية المتعدد الأوجه :

الى جانب حركة البحوث الضخمة التى حركها الاختبار المتعدد الأوجه فقد كان له فى ذاته خصوبة عالية من حيث اشتقاق مقاييس جديدة منه أو تأسيس اختبارات وتطويرها معتمدة عليه .

فمن المقاييس الجديدة الكثيرة التى استخرجت منه عن طريق استعمال مفاتيح تصحيح جديدة نذكر ما يلى :

١ - مقياس التعصب (Pr.) ويهدف الى قياس العوامل السيكولوجية المرتبطة بالتعصب ضد جماعات اقلية .

٢ - مقياس المكانة الاجتماعية - الاقتصادية (St) ويقيس بغض العوامل السيكولوجية المرتبطة بالمكانة الاجتماعية - الاقتصادية للشخص .

٣ - مقياس الانحراف الجنسى .

٤ - مقياس قوة الأنأة .

ويعتبر مقياس قاييلور للقلق الظاهر (٥٩) الذى سبقته الإشارة اليه ، من المقاييس التى اعتمدت على اختبار الشخصية المتعدد الأوجه . وقد استعمل هذا المقياس استعمالاً واسعاً فى البحث وفى الممارسة الاكلينيكية . وقد انشئ أصلاً للاستعمال فى التجارب لاختبار فروض معينة تتصل بتأثير الدافع على التعلم ، وتحقيقاً لهذا اختيرت عبارات معينة من الاختبار المتعدد الأوجه على أنها تدل على تعبير ظاهر عن القلق ، وكانت نتيجة هذا الاختبار ٥٠ عبارة ، زيد عليها عبارات اضافية لا تتصل بالقلق ومختارة من مقاييس الضبط الثلاثة ل ، ف ، ك . وقد ترجمه الى العربية الدكتوران مصطفى فهمى ومحمد أحمد غالى .

كذلك نتج عن الجهود التى بذلت لتعديل الاختبار المتعدد الأوجه بحيث يستعمل مع تلاميذ المرحلة الثانوية والطلبة الجدد فى الكليات ان انشئ

اختبار منسوتا للارشاد النفسى Minnesota Counseling Inventory

ويتكون هذا الاختبار من ٤١٢ عبارة على المفحوص أن يجيب عنها بصواب أو خطأ ، ومعظم العبارات مأخوذة من اختبار الشخصية المتعددة الأوجه . ويعطى اختبار منسوتا للارشاد النفسى درجات فى سبعة مجالات هى العلاقات الأسرية ، والعلاقات الاجتماعية ، والثبات الانفعالى ، والمسايرة ، والتكيف مع الواقع ، والمزاج ، والقيادة . وعلى الرغم من أن مقاييس اختبار الارشاد النفسى تسمى باسم الطرف المقبول ، أى الثبات الانفعالى مثلا بدلا من الطرف الآخر وهو الاضطراب الانفعالى ، فانها تتشابه مع مقاييس الاختبار المتعدد الأوجه المسماة بأسمائها المرضية، فمثلا مقياس المسايرة، فى اختبار الارشاد النفسى يشبه الى حد كبير مقياس الانحراف السيکوباتى (Pd) فى الاختبار المتعدد الأوجه ، وكذلك مقياس التكيف مع الواقعين يشبه كثيرا مقياس الفصام (Sc) ولهذا الاختبار مقياسان للضبط يشبهان مقاييس الضبط فى الاختبار المتعدد الأوجه وتشير بعض الدراسات (فى ٢٧ : ٤٩٦) الى أن مقاييس هذا الاختبار ذات قدرة ايجابية ولكنها ضعيفة على تمييز التلاميذ الذين يعرف انهم غير متوافقين أسريا من غيرهم ممن يتمتعون بتوافق جيد .

وللاختبار ثبات باعادته بعد ثلاثة شهور يتراوح بين ٧٠ ٪ ، سمير .
الا أن تفسير هذا الاختبار يحتاج الى حذر كبير . وقد نقل هذا الاختبار الى اللغة العربية محمد عماد الدين اسماعيل وسيد عبد الحميد مرسى باسم مقياس الارشاد النفسى ، ولا تتوافر عنه بيانات سيكومترية فى البيئة المصرية .

ومن الاختبارات التى اعتمدت على الاختبار المتعدد الأوجه اختبار كاليفورنيا النفسى California Psychological Inventory فهو يشتمل نصف عباراته من الاختبار المتعدد الأوجه ومن أمثلة مقاييسه : السيطرة ، النزعة الاجتماعية ضبط النفس ، الانجاز عن طريق المسايرة ، والانجاز عن طريق الاستقلال والمرونة ، وتعتبر انستازى (١٨ : ٤٤٨) هذا الاختبار من افضل اختبارات الشخصية المستعملة حاليا ، وقد خضع لدراسات وتحسينات عديدة أعطته هذه المكانة . ولهذا الاختبار صورة لتلاميذ الثانوى هى اختبار كاليفورنيا للشخصية California Personality Inventory نقلها

الى اللغة العربية جابر عبد الحميد ويوسف الشيخ ، ولا تتوافر عنها بيانات سيكومترية فى اللغة العربية . كما توجد صورة من هذا الاختبار لتلاميذ

المرحلة الابتدائية نقلها الى اللغة العربية عطية محمود هنا باسم : اختبار الشخصية للأطفال ، (*) .

نماذج أخرى ممثلة من اختبارات الشخصية المعتمدة على التقرير الذاتى

عرضنا بشئ من التفصيل اختبار الشخصية المتعددة الأوجه لأنه « النموذج الأمثل » لاختبارات الشخصية التى تعتمد على التقرير الذاتى للمفحوص ، الا ان الاختبارات التى تنتمى الى هذا النوع كثيرة ومتنوعة الى درجة يكاد يستحيل حصرها جميعا وخاصة فى هذا المجال المحدود ، لهذا نختار نماذج ممثلة من هذه الاختبارات التى لم تناقش فى الصفحات القليلة الماضية مراعين أن يكون لكل منها ميزته التى تجعله جديرا بهذا العرض الموجز .

قائمة موني للمشكلات Mooney Problem Check list وقد وضعها روس موني Ross L. Mooney وزملاؤه . ولها عدة صور من مراحل التعليم المتوسط حتى الجامعة والراشد ، ويحدد فيها المفحوص مشكلاته فى أحد عشر مجالا منها الأخلاق والدين ، النواحي المالية والمعيشية ، والتوافق للعمل المدرسى ، والعلاقات الاجتماعية ، وتشير الدرجة المرتفعة الى الحاجة الى الاشارة النفسى ، وتتخذ العناصر التى يحددها المفحوص أساسا للمناقشة الفردية والجماعية وقد استخدمت هذه القائمة فى عدة دراسات عربية .

اختبار التوافق لـ بل The Bell Adjustment Inventory

منذ صدور هذا الاختبار فى عام ١٩٢٤ (٢٤) وهو يستعمل على نطاق واسع ، والصورة الخاصة بالطالب مصممة للانتقاء السريع لطلاب المدارس الثانوية والكليات بهدف الارشاد النفسى ، وهذه الصورة تعطى

(*) أجرى عطية محمود هنا دراسة على : اختبار الشخصية للأطفال ، استعرض فيها صدقه وثابته فى البيئة المصرية . راجع :
عطية محمود هنا اختبار الشخصية للأطفال وقيمه فى البحوث النفسية . المجلة الاجتماعية القومية مايو ١٩٦٥ .

درجات توافق في مجالات أربعة هي : الأسرة ، الصحة ، الناحية الاجتماعية ، والناحية العاطفية . والصورة الأخرى خاصة بالكبار وهي أقل استعمالاً ، وقد زيد عليها مجال خامس هو مجال التوافق المهني . وقد ترجم الصورة الخاصة بالطلاب محمد عثمان نجاتي . ولا تتوافر عنه بيانات سيكومترية في البيئة المصرية .

مقياس البورت لدراسة السيطرة والخضوع Allport A-S Reaction Study

يعتبر هذا المقياس من أقوى اختبارات الشخصية المبكرة ، ويوصف بأنه يقيس السيطرة ascendance والخضوع submission لأنه يقيس نزعة الشخص إلى أن يسود ويسيطر على زملائه ، أو أن يساد ويسيطر عليه في علاقات المواجهة اليومية . ويبدأ كل عنصر بوصف موجز لموقف من مواقف الحياة اليومية العادية في اجتماع أو في مدرسة أو سيارة عامة ، ويذكر أسلوبين أو أربعة لمواجهة الموقف ، وعلى المفحوص أن يحدد أي الأساليب يعبر أكثر عن سلوكه . وتختلف الاستجابات في درجة السيطرة والخضوع التي تمثلها وتعطى لها أوزان مناسبة لهذه الدرجة في التصحيح . وفضلاً عن شيوع هذا الاختبار فإنه قد أثر في كثير من الاختبارات التي انشئت بعده ومن المهم أن نشير إلى أن السيطرة كانت إحدى السمات التي برزت بوضوح في دراسات التحليل العاملي .

البروفيل الشخصي لجوردن Gordon Personal Profile

يعتبر هذا الاختبار ، وقائمة الشخصية للمؤلف نفسه ، من اختبارات الشخصية التي تعتمد على الاختيار الإجباري forced - choice للتغلب على ما يسميه المرغوبية الاجتماعية ، أي ميل المفحوص إلى أن يبدو مقبولا اجتماعيا ، وكذلك للاقلال من الغموض وبعض الصعوبات الأخرى التي تصاحب عملية تصحيح اختبارات الشخصية .

ويقيس البروفيل أربعة جوانب للشخصية هي : السيطرة ، والمسئولية ، والاتزان الانفعالي ، والاجتماعية (أو النزعة الاجتماعية) وهذه الجوانب

مستقلة نسبيا ، والبروفيل مناسب للاستخدام مع طلاب المدارس الثانوية والجامعة والراشدين .

ويتكون البروفيل من ١٨ مجموعة من العبارات الوصفية تشمل كل مجموعة على أربع عبارات وتمثل كل عبارة إحدى سمات الشخصية الأربع ، المشار إليها من قبل . وتتشابه جملتان من الأربع من حيث أن لها قيمة تفضيلية عالية . أما الجملتان الأرتان فتتساويان في القيمة التفضيلية المنخفضة . ويطلب من المفحوص أن يضع علامة على جملة من الجمل الأربعة على أساس أنها تشبهه إلى حد بعيد وعلى جملة أخرى على أساس أنها تشبهه إلى أقل درجة وبهذا لا يستطيع المفحوص أن يستجيب استجابة قبول لجميع العبارات كما يحدث في كثير من مقاييس التقرير الذاتي الأخرى أي أن البروفيل يعتبر أقل قابلية للتشويه من جانب أولئك الذين يحاولون أن يعطوا صورة طيبة عن أنفسهم .

وقد اقتبس البروفيل وأعدده بالعربية جابر عبد الحميد جابر ، وفؤاد أبو حطب ، ووضعوا له كراسة تعليمات تتضمن وصفا له وتعليمات التصحيح كما وضعوا له معايير مئوية ، وأجريت دراسات عن معاملات المقاييس الأربعة للبروفيل وصدقته ، كما استعمل فؤاد أبو حطب البروفيل في دراسة (١٤) جاءت مدعومة لصدق التكوين الفرضي .

• قائمة لشخصية لجوردون Gordon Personal Inventory

وتعطي هذه القائمة قياسا مريعا ومناسبا لأربع من سمات الشخصية هي : الحرص ، والتفكير الأصيل ، العلاقات الشخصية ، الحيوية . ومن الممكن استخدامها في المدارس الثانوية والجامعة ومع الراشدين . وتستخدم هذه القائمة ، مثل البروفيل الشخصي ، طريقة الاختيار الاجباري ، وهي تتكون من ٢٠ مجموعة من العبارات الوصفية ، وتشتمل كل مجموعة منها على أربع عبارات ، وتمثل كل عبارتين يعتبرهما الأشخاص العاديون متساويتين في قيمة التفضيل العالية ، وعبارتين يعتبرونهما متساويتين في قيمة التفضيل المنخفضة . ويطلب من المفحوص أن يضع علامة أمام عبارة واحدة من العبارات الوصفية على أساس أنها تنطبق عليه أكثر من الصفات الثلاث

الأخرى ، كما يضع علامة أخرى أمام عبارة أخرى باعتبارها أقل انطباقا عليه من العبارتين الأخرتين ، وتتميز القائمة بسهولة التطبيق والتصحيح والتفسير . وقد ترجمها الى العربية فؤاد أبو حطب وجابر عبد الحميد جابر ووضعها كراسة تعليمات تتضمن وصفا لها وتعليمات التطبيق والتصحيح كما وضعت لها معايير مئوية ، وحسبت لها معاملات الثبات والصدق ، واستعملها فؤاد أبو حطب في الدراسة التي سبقت الإشارة اليها (١٤) .

• مقياس المسؤولية الاجتماعية :

تأليف سيد أحمد عثمان ، وهو مقياس ليعند من الأبعاد الاجتماعية - الأخلاقية للشخصية على أساس تصور وتحليل نظري للمسؤولية الاجتماعية متمثلة في عناصرها الثلاثة وهي الاهتمام ، والفهم ، والمشاركة ، وفي المستويات المتدرجة في كل عنصر من هذه العناصر ، كما حددها وناقشها مصمم المقياس (٥) . ويتكون المقياس من صورتين الصورة ت لتلاميذ الثانوى والصورة ك لطلاب الجامعة والراشدين .

وتتكون الصورة ت من المقياس من ١١٥ عنصرا ، وهي عبارات تعكس الرواينا من السلوك أو الآراء ، منها عبارات موجبة وأخرى سالبة ويطلب من المفحوص ان يحدد اجابته وفق مقياس متدرج من اربع نقط .

وتسير الصورة ك من المقياس على النمط ذاته ، وهي مقتبسة ومختارة من الصورة الأولى الا انها تتكون من ٨٥ عنصرا ، منها أيضا عناصر موجبة وأخرى سالبة .

وتتضمن كراسة تعليمات المقياس ، بصورتيه ، شرحا للتحليل النظرى للمسؤولية الاجتماعية ، ثم بيانات عن وصف كل صورة من المقياس ، وعن صدق الصورة ت وثباتها ، وإشارة الى الدراسات التي أجريت على هذه الصورة حتى الآن ، ثم طريقة التصحيح بالنسبة الى الصورتين ، وفي كراسة التعليمات أيضا بيان بإمكانيات الاستعمالات المختلفة للمقياس في صورتيه في النواحي التربوية ، وفي دراسة الشخصية ، وفي العلاج النفسى ، وفي الدراسات الاجتماعية .

والدراسات مستمرة في اعداد معايير المقياس واستكمال دراسات الثبات والصدق وخاصة للصورة ك وقد اجريت طوال العقد المنقضى من منذ ظهور المقياس عدة دراسات عليه أكدت فعاليتها . ويعتبر هذا المقياس أداة بارعة مرجوة تحتاج الى عمل طويل معتد متنوع تزداد معه الأداة دقة وكفاءة ، كما يزداد معه التصور النظري للمسئولية الاجتماعية اتساعا ووضوحا وعمقا .

• اختبارات كاتل في الشخصية :

تعتمد هذه الاختبارات على نتائج البحوث العاملية التي اجراها ريموند كاتل R. Cattell ومعاونوه ، وقد يكون أشهرها اختبار العوامل الستة

عشر في الشخصية .

Sixteen Personality Factor Questionnaire (16 PF).

الذي يقيس ١٦ سمة . كما أعدت قائمة أخرى للمراهقين (من ١٢ - ١٦ سنة) وتسمى Jr. — Sr. High School Personality Questionnaire

وأعدتها باللغة العربية سيد محمد غنيم وعبد السلام عبد الغفار باسم استفتاء الشخصية للمرحلة الاعدادية والثانوية ، وتقيس ١٤ سمة . وكذلك أعدت قائمة ثالثة للمستويات العمرية من ٨ سنوات الى ١٢ سنة وتسمى IPAT Children's Personality Questionnaire وأعدتها باللغة العربية عبد السلام عبد الغفار وسيد محمد غنيم باسم استفتاء الشخصية للمرحلة الأولى ، وقد حسبت المعايير الثابتة للاختبارين الآخرين على عينات مصرية كما حسبت ثبات استفتاء الشخصية للمرحلة الأولى الأولى بطريقة الصورة المتكافئة (ن = ١٩٢ طفلا) ، ولاستفتاء الشخصية للمرحلتين الاعدادية والثانوية بطريقة اعادة الاختبار (ن = ١٢٠ تلميذا) .
الا انه لا تتوافر بيانات سيكومترية عن صدقها في البيئة المصرية .

• قائمة أيزنك للشخصية :

هذا الاستخبار من اعداد عالم النفس الانجليزي أيزنك وتقيس الانبساط مقابل الانطواء ، والعصابية . ونقلها الى اللغة العربية جابر

عبد الحميد ومحمد فخر الاسلام . ولا تتوافر لدينا معلومات سيكومترية عنها في البيئة المصرية .

• استبيان مستوى الطموح :

هذا الاستخبار من تأليف كاميليا عبد الفتاح ، ويتألف من ١٩ سؤالاً مصنفة الى ٧ فئات هي : النظرة للحياة ، الاتجاه نحو التفوق ، تحديد الأهداف والخطة ، والميل الى الكفاح ، وتحمل المسؤولية والاعتماد على النفس ، والثابرة والرضا ، بالوضع الحاضر والايمان بالحظ . وحسبت له المعايير المئينية (ن = ١٢٢ طالبة ، ١٩٤ طالبا) ، وقد ذكرت المؤلفه معاملات ثبات الاستخبار وصدقه وتمييزه بين الأسرياء والعصابيين (*) .

(*) كاميليا عبد الفتاح : كراسة تعليقات استبيان مستوى الطموح ، النهضة المصرية ، ١٩٧١ .

الفصل الرابع عشر

قياس الشخصية عن طريق الاحكام والملاحظات المنظمة

رابنا فى قياس الشخصية عن طريق التقرير الذاتى ان الشخص يختار ويعطى تقديرات يرى انها تعبر عن نفسه ، اما عند قياس الشخصية عن طريق الاحكام Judgments والملاحظات المنظمة systemic observations فان الذى يقوم بتقدير السلوك او الاداء المميز للفرد او المفحوص شخص او اشخاص آخرون ، ويمكن ان نقسم هذا النوع من قياس الشخصية الى قسمين (٩) :

(أ) مقاييس التقدير وتقديرات القياس الاجتماعى .

ratings and sociometric reports

(ب) الملاحظة المنظمة لعينات السلوك :

systematic observation of behavior samples

وسوف نتناول كلا من هذين القسمين بشئ من التفصيل .

(١) مقاييس التقدير وتقديرات القياس الاجتماعى .

يمكن ان تصنف مقاييس التقدير وتقديرات القياس الاجتماعى الى نوعين على اساس من يقومون بهما ، النوع الاول يقوم به الرؤساء او الكبار ، بينما يقوم الافراد او الزملاء بالتقدير والحكم فى النوع الثانى .

١ - اما عن تقديرات الرؤساء او الكبار فنحن نلاحظ ان الرؤساء او المدرسين مثلا يقومون باستمرار بتقديم نوع من الأوصاف لمؤسسيهم او لمن هم فى كفالتهم ، الا اننا لا نستطيع ان نقارن هذه الأوصاف لأنها تختلف فى أسلوبها ، لهذا استعملت مقاييس التقدير حتى يتوافر قدر معقول من التشابه بين التقديرات .

ويتكون مقياس التقدير من قائمة من السمات يجرى عليها التقدير .

وقد يتكون المقياس من قائمة بسيطة من الصلطات يتم الاختيار بينها ، أو قد يكون مقياسا متصلا يتضمن أقساما وصفية متعددة للاختيار منها . . ويوضح المثال الثاني ، (فى ٢٧ : ٥٠٧) هذا النوع من مقاييس التقدير المتصلة :

وقائع مؤيدة	<ul style="list-style-type: none"> * يسعى اليه الآخرون * يحبه الآخرون * يحتمله الآخرون * يتجنبه الآخرون * لا مجال لملاحظته * يسعى الى اضافة أعمال جديدة الى نفسه * يكمل العمل الاضافى المقترح 	١ - كيف تتأثر أنت والآخرون بمظهره وسلوكه ؟
وقائع مؤيدة	<ul style="list-style-type: none"> * يعمل واجباته العادية تلقائيا * يحتاج الى تحريك أحيانا * تحتاج الى تحريك كثير * يعمل واجباته العادية * لا مجال لملاحظته 	٢ - هل يحتاج الى تحريك أو أنه يعمل من تلقاء نفسه ؟

ولكن - نلاحظ أن مقاييس التقدير عن طريق الرؤساء تكمن فيها مصادر متعددة للخطأ منها مثلا ما يسميه كرونباك (٢٧) ، خطأ الكرم ، ويقصد به ميل من يقوم بالتقدير الى اعطاء تقديرات طيبة مؤيدة ، فإذا سئل مدرس مثلا أن يقدر تلميذا ما فيما يختص بالميل الى التعاون ، فإنه سوف يقدر كل التلاميذ الذين لا يثيرون الشغب تقديرا عاليا جدا ، فى أقصى طرف المقياس بينما يهبط بتقدير القلة من التلاميذ مثيرى الشغب . ويكون الارتفاع فى تقدير الغالبية رد فعل لتمييزهم عن مثيرى الشغب . ومن أسباب خطأ الكرم أيضا أن الرئيس المقدر قد يحس أن مما يمس قدرته الرئاسية أن يعترف أن مرؤوسيه ليسوا أكفا ، أو أن الرئيس قد يجد من الأيسر عليه أن يحكم احكاما طيبة على مرؤوسيه اجمالا بدل أن يتملى ويطليل النظر ليميز بينهم تمييزا دقيقا .

ومن مصادر الخطأ أيضا في هذا النوع من التقدير الغموض والابهام، فقد يفسر المقدر سؤالا أو مصطلحا تفسيراً مختلفاً ، فقد يفسر شخص القيادة ، ، مثلا على أنها تركيز السلطة في يد القائد ، والانفراد باتخاذ القرارات ، فعندما يقدر هذا الشخص مفحوصا تتوافق فيه هذه الصفات فإنه يقدره تقديرا عاليا ومخالفا لتقدير شخص آخر للمفحوص ذاته ، لأن المقدر الأخير يفهم القيادة على أنها تشجيع للمرءوسين ومساعدة في اتخاذ قرارات جماعية .

ويعتبر أسلوب الاستجابة Response style عند المقدر مصدرا آخر من مصادر الخطأ في مقاييس التقدير بواسطة الرؤساء أو الكبار ، فمثلا نجد أن مقدار ما يندر أن يستعمل أطراف المقياس في وصف المفحوصين ، بينما يميل شخص آخر إلى الاسراف في استعمال هذه الاطراف ، أى يميل إلى تصنيف المفحوصين إلى أقسام متميزة جدا . ومن ثم نجد أن مثل هذين المقدرين يختلفان عندما يقدران شخصا واحدا .

وما يؤثر في تقديرات الرؤساء أيضا أن معلوماتهم عن المفحوصين قد تكون محدودة ، فقد يختلف تقدير مدرس التربية الفنية ومدرس اللغة العربية في تقدير ردود فعل تلميذ معين للبهزوط ، ويرجع الاختلاف إلى أن كلا منهما يعرف التلميذ ويلاحظه في عدد محدود من المواقف ، وعلى هذا تكون معرفته به محدودة وليست شاملة .

ومن مصادر الخطأ في هذا النوع من التقدير « أثر الهالة » أو « وقع الهالة » halo effect ويحدث هذا الأثر أو الوقع عندما يتأثر المقدر تأثيرا قويا بالانطباع العام عن المفحوص . فيتكون عنده رأى عام أو تقدير عام عن قيمة الشخص يؤثر كثيرا في تقديره لسمات هذا الشخص الخاصة . وهذا النوع من الخطأ يجعل نمط السمات في الفرد غامضا عند المقدر .

الا أنه من الممكن تلافي كثير من مصادر الخطأ هذه إذا راعينا :

١ - الاختيار السليم لمن يقومون بالتقدير ، من حيث قدرتهم ودرجة اندماجهم وتأثرهم شخصيا بتقديرهم للمفحوصين ، فقد سبقت الإشارة إلى

أن المقدر قد يميل في تقديره . حيز لصالح مرؤوسيه لأن في هذا تأكيد
لكفاءة رئاسته .

٢ - كلما زادت معرفة المقدر بالمفحوصين كانت تقديراته أكثر صدقا .

٣ - إذا استعملنا مقدرًا بعينه أكثر من مرة فمن المفيد أن نحفظ
بتقديراته لمقارنتها والتعرف على أسلوب استجابته ، أو ما يمكن أن نسميه
خطأه الثابت .

٤ - من الممكن زيادة ثبات مقاييس التقدير عن طريق الاعتماد على
أكثر من مقدر واحد . على أن يكونوا على معرفة كافية بالمفحوصين .

٥ - مما يزيد من قيمة مقاييس التقدير أن تعد بعناية ودقة . وبصفة
عامة نجد أن المقياس ذا الخمس نقط يعطى تمييزا أكثر دقة من المقياس أو
القائمة التي تعتمد على إجابات نعم أو لا . كما أن مقياس الخمس نقط ينبه
إلى أنواع متباينة من الاختلافات والانحرافات . وكذلك يساعد الاختيار
الاجباري الذي تكلمنا عنه في بعض اختبارات التقرير الذاتي ، على زيادة
قيمة مقياس التقدير . والهدف من الاختيار الاجباري هو فصل عمل المقدر
فيما يتصل بوصف ما يعمله المفحوص عن تقويم هذا العمل ، أي أن المقدر
يكون مسئولاً عن وصف المفحوص ، بينما يقوم المسئول عن اتخاذ القرار
بتقويم هذا النوع من مقاييس التقدير القائم على الاختيار الاجباري أكثر
ملاءمة لقرارات المؤسسات فيما يتصل بالانتقاء والتصنيف ، ولكنها لا تناسب
كثيراً عملية التوجيه التربوي أو الوصف التفصيلي للفرد .

أما عن صدق مقاييس تقدير الرؤساء فتشير الدراسات التي أجريت
عنها إلى ضرورة أخذ هذه المقاييس بحذر فيما يتصل بالحكم على الأفراد من
جهة ، وباستعمالها للتحقق من صدق بعض مقاييس واختبارات الشخصية
من جهة أخرى . ومن أمثلة هذا أن طريقة تقدير الكبار استعملت في دراسة
صدق مقياس المسئولية الاجتماعية . الصورة ت ، الذي تناوله الفصل
السابق ، حيث قام المدرسون بتقدير تلاميذ العينة على مقياس خاص بالمدرسين
وكان معامل الارتباط بين تقدير المدرسين ودرجات التلاميذ على المقياس

٤٦ر - وقد اعتبر هذا العامل مقبولا ، وخاصة اذا أخذنا في الاعتبار ان تقدير المدرس للتلميذ في أى ناحية من نواحي السلوك تتدخل فيه عوامل كثيرة منها على سبيل المثال ، تحصيل التلميذ الدراسي ، ومستواه الاقتصادي الاجتماعي ، ودرجة جاذبية التلميذ الشخصية ، وتفضيل المدرس وتحيزه الشخصي وهكذا ، ، (٢ : ٢٧) .

وقد أثبتت دراسة قام بها بيترز وكاميل (٥٢) ان ارتباط تقديرات الرؤساء لدى معرفة العاملين ودرائتهم بأعمالهم مع هذه المعرفة والدراية كما قيست باختبار خاص ، كان الارتباط بينهما ٠٢٥ر . بينما وصل هذا الارتباط في بعض الأقسام ٠٥٥ر ، كما أظهرت دراسات أخرى ارتباط التقديرات بطول المدة التي يعرف فيها المقدر المفحوصين ، وبمقدار حبه لهم (في ٢٧ : ٥١٨) .

ولكن مقاييس تقدير الرؤساء والكبار ، على الرغم من حدودها وتعتبر مصدرا غنيا بالمعلومات التي لا يمكن الحصول عليها الا عن طريقها .

٢ - والنوع الثاني من مقاييس التقدير هو تقدير الأقران *peer ratings* التي تعطى معلومات تملأ في قيمتها على تقديرات الكبار أو الرؤساء ، وحتى عندما تتوافر تقديرات الرؤساء فان تقديرات الأقران والزملاء تعطى بيانات عن جانب مختلف من الشخصية . فاذا كان رئيس واحد هو الذي يقوم بالتقدير فان قيام مجموعة من الأفراد أو الزملاء بالتقدير سوف يعطي فرصة لعدد كبير من الأحكام مما يترتب عليه ان متوسط التقديرات بالنسبة الى سمّة معينة من سمات الشخصية سوف تكون أكثر ثباتا من التقدير الذي يقوم به رئيس واحد . ومن الممكن ان ننظر الى تقديرات الأفراد على أنها تعبير موضوعي عن سمّة الفرد ، فالسمّة تقوم الى حد ما على السلوك الفعلي ، الا ان نمط العلاقات الاجتماعية في داخل الجماعة لا شك يؤدي الى وجود نوع من التحيز .

ومن الأساليب المستعملة في تقدير الأفراد ما يسمى أسلوب تعيين الأسماء *nomination* أو الترشيح ، ويستعمل هذا الأسلوب كثيرا في القياس الاجتماعي *Sociometry* . وفي هذا الأسلوب يسأل كل عضو من

أعضاء الجماعة ان يسمى عددا محددا من الأشخاص الذين يتميزون بشكل واضح في ناحية من النواحي ، مثل القيادة ، كما يسمى أيضا الأشخاص الذين تنقصهم هذه الخاصية ، وبالنسبة الى الأطفال فان عملية التسمية قد تتخذ صسورة لعبة التخمين ، أو اختبار ، خمن من ، Guess Who الذى يصف الأدوار المختلفة التى يمكن ان يلعبها الأطفال ويسمى كل عضو فى الجماعة أسماء الأبطال الذين يعتقدون ان كل وصف يلائمهم . وترسم لكل طفل صفحة نفسية عن طريق عدد المرات التى اشير اليه فيها لكل وصف .

وإذا كان من الممكن دراسة خصائص الفرد بطريقة خمن من ، فان الرسم الاجتماعى (السوسيوجرام Sociogram) يعطى فهما واستبصارا أعمق فى تمييز التجمعات الداخلية فى الجماعة ، وبيان البناء الهرمى للقيادة والمنعزلين ، أى هذا الرسم الاجتماعى يعتبر طريقة لدراسة البيئة الاجتماعية فى الجماعة (١٦ : ٤٤٣ وما بعدها) .

والذى أنشأه الرسم الاجتماعى وطوره مارينو Moreno (٢٩) . الا ان الأسلوب قد عدل بعدة طرق قللت من فاعليته فى سبيل السهولة واليسر فى الاستعمال . وأفضل طريقة فى استعماله هى ان يسأل أعضاء الجماعة ان يختاروا رفقاء لهم فى نشاط معين . ولا بد ان تكون التعليمات متصلة بأنشطة حقيقية للجماعة ، كما يجب ان تكون الاختبارات حقيقية ويجب ان يعلم أعضاء الجماعة ان المعلومات التى يعطونها سوف تصان سرقتها ، وان البيانات المعطاة سوف تستخدم فيما أخبر أعضاء الجماعة انها سوف تستخدم فيه ، مثلا لاختيار لجان ، أو جماعات عمل فرعية ، أو تنظيم الجلوس فى فصل دراسى وهكذا .

ويعتمد الرسم الاجتماعى الذى نحصل عليه على السؤال الذى توجهه ، فان نمط الرسم الذى ينتج عن سؤال مجموعة من التلاميذ عن اختياراتهم لأقرانهم الذين يحبون ان يستذكروا معهم سوف يختلف عن نمط الرسم الناتج عن سؤال عن اختيارات من يحبون اصطحابهم فى رحلة .

ويشمل مصطلح التقدير بالقياس الاجتماعى Sociometric rating كل طرق تمييز العلاقات الاجتماعية بين أعضاء الجماعة . وليس هنالك

فصل قاطع بين تقدير الأفراد الوصفى ، وبين التقدير بالقياس الاجتماعى ،
الا اننا نجد النوع الأخير بصفة محدودة باستئالة عن يحبهم المقدر اكثر
او يفضل العمل معهم ، ولهذا فهى تتصل باستجابة المقدر كما تتصل
بشخصيته .

ولتقديرات الأقران استعمالات متعددة منها . ان قائد الجماعة
يستعملها فى التعرف على الأشخاص الذين يحتاجون الى عناية خاصة ، او
يستعملها فى اعادة تنظيم الجماعة حتى تؤدي وظيفتها بشكل افضل . كما
ان هذه التقديرات يمكن ان تتخذ اساسا للانتقاء والتصنيف ، وهناك من
يعتبرها انقى مقياس للقيادة ، اما بالنسبة الى المرشد النفسى ، فان تقديرات
الأقران توضح خصائص الشخص التى تعوق قبول الجماعة له ، وخاصة
عندما ينقص الشخص الاستبصار الكافى فى رأى الآخرين فيه .

ويذهب كثير من الباحثين الى ان تقديرات الأقران اثبتت انها احدى
وسائل التقدير التى يمكن ان يوثق بها الى درجة بعيدة لعدة أسباب هى :

١ - ان عدد المقدرين كبير ، وقد اشرنا منذ قليل الى أهمية هذا فى
زيادة ثبات التقديرات بصفة عامة .

٢ - ان اقران الشخص عادة ما يكونون اكثر قدرة من غيرهم على
ملاحظة السلوك المميز له ، وهم من ثم افضل فى الحكم على سمات معينة
تتصل بالعلاقات الاجتماعية من المدرسين او المشرفين او أى ملاحظين
خارجين .

٣ - ان آراء أعضاء الجماعة الآخرين ، صوابا كانت ام خطأ ، تؤثر
فى سلوكهم وتصرفاتهم وبالتالي فان هذه الآراء تحدد جزئيا طبيعة تفاعلات
الشخص مع الجماعة . فمثلا ، اذا كان أعضاء جماعة يرون ان أحد أعضاء
جماعتهم ، وليكن اسمه عليا مثلا ، رجل يمكن الاعتماد عليه والثقة به ،
فان هذا الرأى يوجه ويحدد نوع تعاملهم مع على ، وغالبا ما يحدد هذا
العامل نوع السلوك من جهة على كما يحدد نوع التفاعل من جهة الجماعة
فالجماعة تعامله على انه ثقة ، وهو يسلك تجاه الجماعة وفق توقعاته منه

(٥ : ٩٨ - ١١٠) ولا شك أن هذا يعطى تقديرات الأقران صفة أقرب الى الواقعية .

(ب) الملاحظة المنظمة لعينات السلوك

من الممكن أن تعطينا الملاحظات المنظمة وصفا أكثر دقة للسلوك أو الأداء المميز إذا قورنت بأساليب التقرير الذاتى أو الأحكام والتقديرات التى يصدرها الآخرون عن الشخص ، ذلك لأنه فى التقرير الذاتى من المحتمل أن يتدخل التذكر الانتقائى ، بل انه يحدث أيضا عندما تستعمل أساليب التقدير والحكم المختلفة . وكذلك نجد أن الملاحظة فى هذه الأساليب الأخيرة ليست قائمة على الاختيار المنظم للمواقف التى يلاحظ فيها السلوك ، بل انها تقوم على ملاحظة المواقف كيفما اتفق وقوعها ، وفى هذه الأساليب أيضا لا يلاحظ الشخص فى كل المواقف . لهذا قلنا ان الملاحظة المنظمة تعطى وصفا أدق للسلوك المميز .

ومن السهل أن نلاحظ سلوك مفحوص ما فى ظروفه العادية . وهذا ما يسمى بالملاحظة الميدانية . هذا النوع من الملاحظة يكون أكثر ملاءمة فى بعض الدراسات من الملاحظات المقننة ، فكثير من الباحثين يرى أنه من المستحيل أن نفهم الشخصية الا اذا لاحظنا المفحوص وهو يستجيب الى الظروف التى لها مغزى بالنسبة اليه ، لأن المثيرات المختلفة لها معان ومغاز مختلفة بالنسبة الى الأفراد المختلفين . وعلى هذا فهم يرون أن المواقف المقننة حيث تلاحظ ردود أفعال الأشخاص لموقف بعينه ، تبرز أنماط السلوك الهامة كما تبرزها الظروف العادية ، غير المتشابهة ، التى يعيش فيها الفرد . الا ان المشكلة هنا تكمن فى كيف نرى قدرا كافيا من سلوك الفرد ، والحصول على تسجيلات لملاحظات السلوك يمكن الاعتماد عليها . أى أن المشكلة هنا ذات شقين : الأول ، عينة السلوك وما يرتبط بها من صعوبات ، والثانى ، عيوب أو أخطاء من يقرم بالتسجيل . ولنتكلم عن كل منهما بشئ من التفصيل .

اختيار عينة السلوك :

لكى تعرف السلوك أو الأداء المميز لشخص ما لابد أن نعرف ما يميز تصرفه فى موقف بعينه، ولكن المواقف تتغير من يوم الى يوم ومن لحظة الى

أخرى ، والطريقة الوحيدة لنطمئن الى أن ما نلاحظه هو السلوك المميز هي أن ندرس هذا الشخص في مواقف ومناسبات عديدة ، وهو أمر مجهد ومكلف . فالواقع أنه لابد أن نصل الى نوع من التوفيق بين الاختيار الدقيق لعينة السلوك والاقتصاد في الانفاق والجهد .

وانه من العسير جدا أن نقوم باستنتاجات عن فروق بين الأفراد من ملاحظة سلوكهم في مواقف معينة ، لأنه لا يمكن ملاحظة فردين في موقف واحد بعينه ، وحتى إذا كان الموقف واحد في ظاهره فإن الظروف السابقة الخاصة بكل منهما تجعلهما يسلكان سلوكا مختلفا . ولهذا فإن أقصى ما يمكن أن تقدمه لنا الملاحظات المقارنة هو بيان كيف يسلك أشخاص مختلفون في ظروفهم الحالية ، ولكن هذه الملاحظات لا تضمن لنا أن تستمر هذه الفروق عندما تتغير طبيعة هذه الظروف .

ومن أفضل الأساليب التي تؤدي الى المقارنة الدقيقة ما يسمى بالعينات الموقوتة time sampling وفي هذا الأسلوب يخطط مسبقا جدول أو قائمة للملاحظات ويصمم هذا الجدول بشكل عشوائي يضمن أن يلاحظ كل مفحوص في ظروف متشابهة مع الآخرين بحيث تمكن المقارنة . فمثلا عند دراسة الاتصالات الاجتماعية لأطفال ما قبل المدرسة يوضع جدول الملاحظات لمدة دقيقة واحدة ، وبعد ملاحظة الطفل لمدة دقيقة يكتب الملاحظ تقريراً كاملاً عما لاحظته من تفاعلات الطفل الاجتماعية . ويلاحظ الأطفال وفق نظام مرسوم ويعدل من يوم الى يوم ، وفي خلال الدراسة كان كل طفل يلاحظ عدد مرات متساوية في خلال الخمس دقائق الأولى من ساعات اللعب الحر ، ثم الخمس دقائق الثانية وهكذا . والعينات الموقوتة قصيرة الأجل والموزعة توزيعاً حسناً عدة مزايا ، فالصورة المتراكمة من الملاحظات فيها تكون أكثر تمثيلاً للسلوك من نفس القدر من المعلومات التي يحصل عليها من ملاحظات من مرات أقل وذات فترات طويلة ، كما أن أخطاء الذاكرة تقل هنا لأن الملاحظ يقوم بالتسجيل أثناء وبعد الملاحظة مباشرة . والعينات الموقوتة أكثر مناسبة لتسجيل حقائق معينة يمكن أن تصاغ في صورة رقمية ، مثل عدد مرات الاتصال الاجتماعي بالآخرين .

ومن الممكن الحصول على عينة أكثر اتساعاً بأسلوب « السجل اليومي »

day record الذى انشاء باركر وزملاؤه (٢١) . كان هدفهم هو دراسة نمط حياة الطفل ومع الاهتمام الخاص بالمجالات التى يتحرك فيها .

وللسجل اليومى مزايا ان يعطى صور عن تتابع النشاط ، ولكن له عيوبه ، فاذا كان الملاحظ معروفا ومكشوفاً تأثر السلوك الذى يلاحظه ، والاعتماد على يوم كامل بعينه لا يكفى للحصول على معلومات مميزة كاملة عن المفحوص .

ومن المهم ان نلاحظ بالنسبة الى عينة السلوك الملاحظة عن طريق العينة الموقوتة ما يلى :

١ - ان الاستجابات التى يعدها الملاحظ متشابهة او واحدة قد يكون لها فى الحقيقة معان مختلفة . وان المواقف التى تبدو له متشابهة قد تستدعى استجابات متباينة .

٢ - ان الاستنتاجات المشتقة من موقف واحد ، حتى ولو كانت على اساس تراكم ملاحظات ، تكون ، اى الاستنتاجات صحيحة بالنسبة الى هذا الموقف فقط .

٣ - ان الملاحظات المطلوبة والمناسبة للحصول على معلومات ثابتة ويمكن الثقة بها ، يعتمد هذا العدد على المشكلة ، الا انه بصفة عامة نجد ان عددا كبيرا من الملاحظات القصيرة افضل من عدد قليل من الملاحظات التى تعتمد على فترات طويلة .

عيوب من يقوم بالملاحظة :

الواقع ان بعض عيوب الملاحظين قد يرجع الى طبيعة المواقف الاجتماعية ذاتها التى تكون موضوع ملاحظة حيث تحول درجة تعقدها الكبيرة دون التقرير الشامل للمستغرق لها . اما العيوب الأخرى فترجع الى الملاحظ نفسه .

واذا كانت أخطاء الملاحظ مجرد حذف أو اغفال عشوائى لبعض ما

يلاحظ فليس لهذا وزن كبير ، واما ما يجب الاهتمام به فهو الأخطاء المنظمة ،
كان يبالغ باستمرار وبصفة مطردة في تأكيد بعض أنماط الأحداث ، كما يفغل
الإشارة الى بعضها الآخر .

عندما يلاحظ مراقبون موقفا واحدا فانهم يعطون تقارير متباينة بدرجة
كبيرة ، ذلك لأن كل ملاحظ عنده حساسية لأنماط معينة من السلوك أكثر من
غيرها والوضع الأمثل هو ان يبنى الملاحظ انطباعه على كل تصرف يمكن ان
يكشف أو يعكس خصائص أساسية ، ولكن عندما يركز من يقوم بالملاحظة
والمراقبة على شيء بعينه فانه بالضرورة لابد ان يفغل ملاحظة أشياء أخرى .

وثمة عيب آخر وهو أن الملاحظين يفسرون ما يرون ، فاذا قام الملاحظون
بتسجيل الحقائق الموضوعية ، فقد يصل آخرون الى تفسيرات مختلفة تماما ،
ولكن الناس يميلون الى تفسير ما يرون ، وعندما يفسرون فانهم يميلون الى
تجاهل الحقائق التي لا تتفق مع تفسيرهم بل انهم قد يخترعون حقائق
لتكمل تفسيرهم .

ومما يقلل من اثار هذه العيوب أن يتم التسجيل بشكل منظم ، وان
يكون كلما أمكن ، في صورة وحدات سلوكية يمكن عدها ، فمثلا من الممكن
ان نصف مدى انتظام عمال مصنع في عملهم عن طريق التسجيل الزمني
للحظات التي يكونون فيها في عمل فعلي ، واللمظات التي يتشغلون فيها عن العمل
للحصول على الأدوات ، أو زيارات زملائهم ، أو يتعرضون فيها لمؤثرات
التشتيت . وحتى اذا كان السلوك الملاحظ على درجة كبيرة من التنوع بحيث
يصعب وضعه في صورة وحدات سلوكية فمن الممكن ان تحدد أقسام السلوك
وتعد قبل ويكون عمل الملاحظ هو أن يصنف السلوك الذي يلاحظه وفق هذا
التقسيم ، ويضع علامة في القسم المناسب لكل أسلوب يلاحظه كما هو الحال
في أقسام بيلز Bales (راجع ٦) .

وعلى الرغم من فائدة التسجيل والعد الموضوعي لوحدات السلوك
وملائمته للبحوث العلمية فان فائدته محدودة في التوجيه والارشاد النفسي .

اما السجلات القصصية أو الوقائية anecdotal records فانها .

وان كانت تنقصها دقة الموضوعية ، فانها تعطى وصفا حيا للشخص ، لأن الملاحظ فيها يوجه اهتمامه الى أى سلوك يبدو جدير بالملاحظة بدل أن يركز على سمات معينة للمفحوصين كلهم الا ان الملاحظ فى السجل القصصى يصف بدقة ما يلاحظه فاصلا بين الوقائع وتفسيراتها . وعندما تتراكم هذه السجلات القصصية فانها تعطى صورة اكثر غنى من أى وسيلة أخرى على هذه الدرجة من البساطة . ومسئولية الملاحظ الكبرى هنا تتركز فى اختيار الوقائع التى تستحق التقرير . وفى ان يتحرى الموضوعية بقدر الامكان .

الفصل الخامس عشر

قياس الشخصية عن طريق الأداء والأساليب الاسقاطية

على الرغم من أن اختبارات الأداء والأساليب الاسقاطية تستعمل في قياس الشخصية منذ أكثر من أربعين عاما فإنها لم تصل الى درجة النضج التي وصلت اليها الاختبارات والأساليب والطرق التي تكلمنا عنها حتى الآن . لقد كان الحلم الأكبر للمشتغلين بقياس الشخصية أن يصلوا الى وسائل قياس • تمثل السلوك مباشرة وليس انطباعات متميزة أو تقارير ذاتية غير دقيقة ، ووسائل قياس تعطى نتائج كمية ، وتسمح بالمقارنة المباشرة بين الأفراد في المواقف ذاتها • وكانت اختبارات الأداء والوسائل الاسقاطية - وليس بينهما تمييز قاطع كما يرى كرونباك (٢٧ : ٥٢٩) - كانت ضمن هذه الوسائل ، وإن كانت لم تحقق ، كما سبق ، نضجا يعادل ما وصلت اليه أساليب قياس الشخصية الأخرى •

واختبارات الأداء التي تهدف الى قياس الشخصية تصمم بحيث تؤدي الى ظهور دلائل ومؤشرات وعينات من السلوك المميز حقا للشخص ، ويمكن أن نحدد الملامح العامة لمثل هذه الاختبارات فيما يلي :

١ - أن يكون الموقف المثير موحدا بقدر الامكان بالنسبة الى المفحوصين كلهم •

٢ - يصمم الموقف بحيث يسمح بظهور تنوع في انماط السلوك التي يهدف المختبر الى ملاحظتها •

٣ - يوجه المفحوص الى الاعتقاد أن خاصية معينة هي التي تقاس بينما يقوم الملاحظ بملاحظة خاصة أخرى أو جانب آخر من السلوك أو الأداء •

٤ - يسجل الملاحظ بدقة طريقة المفحوص في الأداء وليس كمية السلوك المؤدى •

والمزية الكبرى للملاحظة في اختبارات الأداء أنها تكشف عن خصائص لا تظهر الا قليلا جدا في أنشطة الشخص العادية من هذه الخصائص مثلا الشجاعة ، ورد الفعل للحبوط ، وعدم الأمانة . ما أن من مزايا هذا النوع من الملاحظة في مواقف الأداء ان رغبة المبحوص في أن يعطى انطبعا حسنا لا تؤثر في مدى صدق الاختبار ، لأن الشخص بسبب هذه الرغبة سوف يكشف عن خصائصه بدرجة أكبر مما لو لم تكن عنده هذه الرغبة ، وعلينا ان نأخذ في الاعتبار هذا الدافع عندما نفسر نتائج ملاحظتنا . ومن مزايا اختبارات الأداء أيضا أنها تسمح الى درجة بعيدة بمقارنة الأشخاص في تعرضهم لظروف واقعية موحدة ومتشابهة .

وتتراوح المواقف التي تستخدم الأداء بين المواقف البنية والمحددة في بنيتها highly structured والمواقف التي لا تتميز ببنية واضحة محددة Totally unstructured . ويقال ان موقفا ما اذا بنية كاملة ومحددة اذا كان له معنى واحد معين لجميع المبحوصين . بينما يكون الموقف الذي لا تتوفر فيه بنية واضحة محددة متضمنا لعدد قليل جدا من الاشارات، او ليس فيه سوى نمط أو نظام ضئيل جدا الى درجة تجعل من الممكن للمبحوص ان يعطيه أي معنى يشاء . ومن أبسط أمثلة المواقف غير محددة البنية الصوت الغريب الذي ينبعث من مكان في المنزل مغلقا يستأثر الظلام صباحا على أمواج السكينة في وسط الليل . . . تسمعه فلا تعرف بالتحديد ما هو ؟ مصدره ؟ يعطيه كل من يسمعه تفسيراً خاصا ، هذا التفسير الخاص يتأثر الى درجة بعيدة بمخاوفنا ، شعورية كانت أم لاشعورية ، وباهتماماتنا ومعرفتنا ، أي أنه في الموقف محدد البنية يعرف الشخص بالضبط ما يتوقع منه ، وكيف يقوم به ، أما في الموقف غير محدد البنية ، فإن الشخص يواجه نفسه ويقودها في الاستجابة له ، وكلما زاد غموض الموقف ، أي زادت بنيته عدم تحديد زادت الفرصة لتعطي الشخص تفسيراً أو أداء خاصة مميزاً له .

وتعتبر المواقف محددة البنية أكثر مناسبة لقياس القدرة ، أو قياس المهي الأداء ، لأنها تجبر كل مبحوص على أن يحاول أمراً بعينه ، وبينما نجد الأساليب غير محددة البنية ، أو الأساليب الإسقاطية ، تستعمل ، بعكس مقاييس القدرة ، مثيرات لا تتميز غالباً بأية بنية محددة .

وفيما يلي نتناول نوعين من اختبارات الأداء لقياس الشخصية ، النوع الأول هو اختبارات المواقف ذات البنية المحددة ، والنوع الآخر هو اختبارات المواقف ذات البنية الأقل تحديدا أو الاختبارات الإسقاطية .

أولا : اختبارات الأداء ذات المواقف محددة البنية :

تدخل ضمن هذا النوع من الاختبارات : اختبارات الخلق والاصرار العمل ، كذلك الاختبارات التي تقيس الأساليب الإدراكية والمعرفية .
perceptual and cognitive styles

لقد كان استعمال تربية الخلق The Character Education Inquiry

الذي وضعه هارتشون وماي Hartshorne and May من أكثر الجهود التي بذلت لتقويم الشخصية عن طريق أساليب كمية وموضوعية خالصة ، وكان اعتقادهما أن سمات الخلق يمكن أن تقاس بدقة عن طريق تعريض الشخص للاغراء في مواقف يظن أنه يمكن فيها أن يخالف المعايير والقواعد الأخلاقية دون أن يكتشف أمره . وكانت السمات التي درسها هذان الباحثان تتضمن ، الصدق ، والأمانة فيما يتصل بالنقود ، والاصرار والداب في العمل ، والتعاون ، الكرم ، (راجع ٢٢ ، ٢٣ ، ٢٤) .

أما اختبارات الأفراد في الأساليب والطسرق التي يتبعونها في ادراك وتناول المشكلات . والواقع أن الدارس الذي يقوم بملاحظة أي سلوك أثناء حل مشكلة يمكن أن يميز بسرعة الاختلافات في الطرق التي يتبعها المفحوصون في مواجهة المشكلات وتخطي العقبات . وتعتبر الاختبارات الفردية ، مثل اختبار وكسلر الذي سبق تناوله ، فرصة لملاحظة هذه الأساليب أو العادات المختلفة بين الأفراد . والمعلومات التي نجمعها عن هذه الأساليب الخاصة تساعد في التنبؤ بنوع المشكلات التي يستطيع الشخص أن يتعامل معها بنجاح ، وأي الأخطاء يمكن أن يقع فيها أكثر ، وقد تفتح أمامنا طريقا لعلاج مناسب يؤدي إلى تحسين الأداء في مواجهة المشكلات .

وتهتم اختبارات العمليات المعرفية أساسا بالطسريقة التي ينظم بها

الشخص المعلومات التي ترد اليه ، هذا التنظيم Organising أو اعطاء نمط patterning كان يمثل الاهتمام الرئيسى عند اصحاب نظرية الجشتالت ، ومعظم الاختبارات التي وضعت فى هذا المجال وضعها باحثون ينتمون الى مدرسة الجشتالت . وقد استعمل هؤلاء الباحثون مثيرات بسيطة لاختبار الجهاز الادراكى كله معتقدين ان المخ جهاز لتحويل المعلومات وانه عن طريق مثيرات او اختبارات ادراكية بسيطة يمكن ان نتعرف على خصائصه المميزة . من هذه المثيرات الادراكية البسيطة اندماج الخفق flicker fusion والحركة الظاهرة apparant movement فعندما ينظر الشخص الى ضوء من خلال جهاز دوار يجيب الضوء ثم يكشفه ، فانه يدرك الخفق عندما يكسبون الدوران ذا معدلات بطيئة ، وكلما زاد معدل الدوران تدريجيا فان الشخص يستطيع ان يقرر النقطة التي يختفى عنها ادراكه للخفق ، وقد وجدت فروق فردية واسعة فى هذه الناحية ، لك من الدلائل ما يشير الى ان نقطة الاندماج هذه تدل على قدرة الجهاز حصى على تسجيل تفاصيل المثيرات الواردة ، وقد استعملت مقاييس اندماج الخفق بنجاح فى تشخيص اصابات المخ .

اما عن الحركة الظاهرة فالمقصود بها انه عندما يعرض ضوء ان جنبا الى جنب مع اطفاء احدهما واناارة الآخر . فتابع سريع فانه يظهر ان الضوء يتحرك قافزا جيئة وزهابا ، وفى الحركة الظاهرة هذه ، وتسمى ايضا ظاهرة فاي ، نجد ان المخ يقوم بتكامل المثيرات فى نمط . وقد درس كلين وشيليزنجر (٢٨) اثر تغيير الفترات بين الاضاءتين ، ووجدوا انه كلما قصرت هذه الفترة كانت هناك عتبة فارقة لبدء ظهور الحركة الظاهرة وعتبة فارقة ثانية تختفى عندها ، اى يرى الشخص بعدها ضواين ثابتين . وقد كان الفاصل بين هاتين العتبتين ، اى مدى الفترات التي تسمح بظهور انطباع الحركة ، اكثر اتساعا عند بعض الافراد عنه عند غيرهم .

ومن اختبارات الأساليب المعرفية ما يهدف الى قياس المرونة العقلية والجمود أو التصلب العقلى . فمما يلاحظ كثيرا ان الذين يفشلون فى حل المشكلات يتميزون بالتشبت والوقوف عند الأفكار الخاطئة ، فهم مثلا يكررون الدخول فى ذات المر المسدود فى المتسامة . والواقع ان المواجهة السليمة للمشكلة تتطلب إعادة تنظيم المجالات الادراكية على اساس المعلومات

أو المتطلبات الجديدة . وقد أنشئت اختبارات لقياس درجة المرونة الإدراكية والعقلية ، نشير منها الى اختبار التاهب Einstellung Test (أو اختبار التوجيه orientation أو اختبار الاتجاه العقلي mental set) الذى وضعه لوشينتز (٤٤) وهو يعرف باسم اختبار اناء الماء . ويستعمل مشكلات توزيع كميات من المياه فى انية محددة الاتساع . ويجرى الاختبار هكذا مثلا ، اذا كان لديك اناء سعته ٧ ارباع جالون وآخر سعته ٤ ارباع جالون فكيف تحصل بالتحديد على ١٠ ارباع جالون ؟ ، وتمشيا مع منطق الاختبار يسمى الاناء ٧ (١) والاناء ٤ (ب) ويكون الحل هو : املا ب تاركا ثلاثة ارباع جالونات فى ١ ، افرغ ب ، املا ب من ١ ، ثم املا ١ . وعلى هذا يحصل على ثلاثة ارباع جالونات حسب القاعدة ١ - ب ، وتحصل على ٧ ارباع جالونات الأخرى من ١ ، أى ان (١ - ب) + ١ = ١٠ . وتكون سلسلة المشكلات كما يلى :

الآنية	المطلوب				
	ب	ج			
مثال ٧	٤	-	١٠	١٢ - ب	
١ - ٢١	١٢٧	٣	١٠٠	ب - ١ - ٢ ج	
٢ - ١٤	١٦٣	٢٥	٩٩	ب - ١ - ٢ ج	
٣ - ١٨	٤٣	١٠	٢١	ب - ١ - ٢ ج	
٤ - ٩	٤٢	٦	٢١	ب - ١ - ٢ ج	
٥ - ٢٠	٥٩	٤	٢١	ب - ١ - ٢ ج	
٦ - ٢٣	٤٩	٣	٢٠	ب - ١ - ٢ ج أو ١ - ج اختبار	
٧ - ١٥	٣٩	٣	١٨	ب - ١ - ٢ ج أو ١ + ج اختبار	
٨ - ٢٨	٧٦	٣	٢٥	١ - ج ا انطفاء	
٩ - ١٨	٤٨	٤	٢٢	ب - ١ - ٢ ج أو ١ + ج اختبار	
١٠ - ١٤	٣٦	٨	٦	ب - ١ - ٢ ج أو ١ - ج اختبار	

ومن الممكن مساعدة الفحوص فى حل المشكلات الأولى . وتؤدى المشكلات من ١ حتى ٥ التى تحل عن طريق تطبيق قاعدة معينة الى تكوين اتجاه عقلى أو تاهب لاستعمال هذه القاعدة أو المعادلة ، بعد ذلك تقدم

مشكلات الاختبار critical والانطفاء extinction . وفى مشكلة الاختبار كما فى المشكلة رقم ٦ مثلا يمكن ان تؤدى القاعدة او المعادلة الى الحل ، ولكن هناك طريقة اسهل للحصول على الاجابة . اما فى مشكلة الانطفاء كما فى رقم ٨ مثلا . فان التأهب ، أى استعمال القاعدة لا يؤدى الى الحل ولكن من الممكن بالنسبة الى المفحوص الذى يتمتع بمرونة أى يكتشف قاعدة اكثر بساطة ولكى يحصل المفحوص على درجة مرتفعة على اختبار التأهب هذا فان عليه ان يتنبه للمشكلة المباشرة امامه ، مستبعدا أى ذكريات عن الحلول السابقة . ومن المحتمل يكون السلوك غير المرن ، المتصلب ، الجامد دليلا على عدم القدرة على الفصل بين المعلومات المتعارضة .

ومن الأمثلة الأخرى لاختبارات المرونة الإدراكية اختبارات الأشكال المضمنة Embedded Figures test حيث يقدم فيها شكل هندسى غريب ثم يطلب من المفحوص استخراجها من مجال أكثر تعقيدا ، وتكون الدرجة هى الوقت الذى يستغرقه المفحوص فى حل مثل هذه المشكلات . وقد أعده للاستخدام فى اللغة العربية سليمان الخضرى الشيخ وأنور محمد الشرقاوى .

ثانيا : اختبارات المواقف غير محدد البنية

ينقسم هذا النوع من الاختبارات الى ثلاثة اقسام :

الأول ، اختبارات حل المشكلات ، والثانى ، الاختبارات الإدراكية غير محددة البنية ، والأخير الاختبارات اللفظية غير محددة البنية . ولنتكلم عن كل قسم من هذه الأقسام .

(١) اختبارات حل المشكلات : من الممكن ان نلاحظ فى اختبار مثل اختبار تصميم المكعبات وبنائها الطريقة التى يتناول بها المفحوص المشكلة او ما يسمى بطريقة المواجهة method of attack كما يمكن ان نلاحظ رد الفعل للحبوط response to frustration عند هذا المفحوص كلما واجهته صعوبة او فشل فى الوصول الى حل لما امامه من مشكلات . الا اننا نستطيع ان نحصل على معلومات افضل اذا عدلنا الاختبار ، واذا حددنا بالضبط ما نريد ملاحظته .

وقد حدد جولدنر (٢٠) Goldner طريقة المواجهة في ضوء متغيرين أكثر تحديدا هما : اتجاه الكل - الجزء whole-part approach والتصلب rigidity وقد استعمل ستة اختبارات منها : اختبار مكعبات معدل ، واختبارات تكوين كلمات من عدد معين من الحروف ، واختبار البرورشاخ . واختبار الوظيفة ، والذي يسأل فيه المفحوص أن يعطى استعمالات أو وظائف مختلفة لأشياء معينة مثل الصندوق أو المكينة أو الورق ، وقد أنشا جولدنر قواعد تصحيح لكل اختبار ، فمثلا بالنسبة الى اختبار الوظيفة تعطى درجة الكل - الجزء حسب ما اذا كانت الاجابة تشير الى استعمال الشيء كله ، مثلا وضع أشياء في الصندوق ، أو كانت الاستجابة تفتت الشيء الى أجزاء كأن تكون الاجابة هي مثلا ، استعمال الصندوق في اشعال النار .

ويقىس جولدنر التصلب أيضا ، فمثلا في اختبار المكعبات يعين التصلب (١) اذا وجد المفحوص صعوبة في الحكم على العدد الصحيح من المكعبات ، و (ب) اذا احتفظ بذات طريقة المواجهة بعد فشلها . أو (ج) اذا يش وتوقف دون أن يكمل حل مشكلة كما يميز التصلب بميل المفحوص الى اعطاء عدد كبير من الاستعمالات والوظائف المتشابهة منطقيا ، بينما تتميز المرونة بالتنوع في الاستعمالات والوظائف .

واذا قارنا أسلوب جولدنر هذا باختبار اناء الماء الذي سبقت الاشارة اليه نلاحظ أنه في اختبار اناء الماء نحصل على درجة تمثل خاصية يمكن حصرها ، الجمهور مثلا ، بينما في بطارية جولدنر تمثل كل درجة تقديرا لعملية عقلية مفترضة ويمكن أن تظهر في الأداء بحيث يمكن استنتاج العمليات العقلية منه . وقد وجد جولدنر تأييدا كثيرا من دراساته لفرضه الذي يذهب الى عمومية السسعتين اللتين لاحظهما اولاهما : اتجاه الكل الجزء ، التصلب .

ومن اختبارات المشكلات التي تستعمل في قياس الشخصية اختبارات تكوين المدرك Concept-Formation فقد استعملت في قياس الشذوذ في العمليات العقلية . ويمكن ان تأخذ اختبار هانفمان - كازانين Hanfmann Kasanin نمونجا لهذه الاختبارات . يستعمل في هذا الاختبار عدد

٢٢ مكعبا خشبيا ذات اللون واشكال وارتفاعات وأحجام مختلفة ، توضع على طاولة ، ومطبوع على الجانب المختص من كل مكعب مقطع عديم المعنى ، وكل مقطع يميز نمطا من المكعبات ، مثلا كل المكعبات التي تحمل مقطع م ن ر تكون صغيرة وطويلة ، ويخبر الممتحن المفحوص بأن المكعبات أنواع أربعة ، نوع ، منها يسمى م ن ر ، وأن عليه ان يكتشف أساس التقسيم ويصنف المكعبات ، ويعمل المفحوص على اكتشاف التصنيف كما يجب ، الا انه لا يسمح له بقلب المكعبات ليرى الأسماء ، ويعد كل محاولة للتصنيف بوضع الممتحن خطأ واحدا ثم يطلب من المفحوص ان يقوم بمحاولة أخرى ، ويستمر هذا الى ان يتم حل المشكلة ، ويعطى المبدأ الذى يقوم عليه التصنيف . وتوضيح ملاحظة المفحوص ما اذا كان يستعمل فروضا منطقية ، مثلا ، ربما كانت كل م ن ر مستطيلات ، أو يستعمل التخمين ، أو يستفيد من التصحيح أو القدرة على استيعاب تأهب غير سليم والتخلص منه وتكوين مدرك جديد . كما يمكن ملاحظة أى أساليب أو طرق شاذة يتبعها المفحوص .

(ب) الاختبارات الإدراكية غير محددة البنية : ويمكن ان نطلق على هذا النوع من الاختبارات الإدراكية ، بصفة عامة ، الاختبارات الإسقاطية أو الأساليب الإسقاطية Projective Technique فى قياس الشخصية . والمميز الرئيسى لهذه الأساليب هى انها تكلف المفحوص بعمل غير محدد البنية نسبيا ، أى عمل يسمح بعدد غير محدد تقريبا من الاستجابات الممكنة ، لهذا فان التعليمات فى مثل هذا النوع من الاختبارات تكون عامة بحيث تترك لخيال المفحوص العنان ، وكذلك فان مثيرات الاختبار نفسه عادة ما تكون غامضة وغير محددة . والافتراض الكامن وراء هذه الأساليب ان الطريقة التى يدرك عليها المفحوص مواد الاختبار ويفسرها ، أى طريقة بنائه للموقف سوف تعكس الجوانب الأساسية لتكوينه النفسى . أى ان مواد الاختبار سوف تعمل فى هذه الحالة كأنها شاشة عرض يسقط عليها الشخص آراءه واتجاهاته وأطماعه ومخاوفه وصراعاته وعدوانيته وهكذا .

وتتميز الأدوات الإسقاطية أيضا بأنها تمثل طرق اختبار مخفاة diseuse لأنه من النادر ان يعلم المفحوص أى نوع من التفسير النفسى سوف يعطى لاستجاباته، كما تتميز هذه الأدوات الإسقاطية بأنها تتخذ اتجاها

شاملا . global . في تقدير الشخصية ، ذلك لأنها تركز على الصورة الكاملة للشخصية كلها وليس على قياس سمات منفصلة .

وفيما يلي نتعرض باختصار لأبرز الاختبارات الإدراكية غير محددة البنية .

١ - اختبار بندر - جشتالت Bender-Gestalt :

وقد وضعته لوريتا بندر في عام ١٩٢٨ ونشرته تحت اسم (اختبار الجشتالت البصري الحركي واستعمالاته الاكلينيكية) ، ويشترك هذا الاختبار مع بعض الاختبارات التي سبقت الإشارة إليها في أنه نبع من أبحاث مدرسة الجشتالت في الإدراك ونلاحظ أن الاختبارات الإدراكية بصفة عامة تقوم على اختبار كيف يتعرض الشخص للمعلومات ويعيد تنظيمها ثم يذكرها ويقررها . وفي اختبار بندر تعرض أمام المفحوص أشكال ذات أنماط مختلفة من التنظيم ، نقاط ودوائر صغيرة وخطوط متعرجة متقاطعة وغيرها ، ويسأل المختبر المفحوص أن ينقل مجموعة الأشكال ، ثم يلاحظ طريقته في الأداء ومدى نجاحه في نقل هذه الأشكال .

وإذا كانت اختبارات الشخصية ذات البناء المحدد مصممة بحيث تقيس سمة معينة فإن اختبار بندر لا يمكن أن تميزه على أساس أنه مقياس لسمة واحدة أو لمجموعة من السمات . ذلك لأن استجابات الأفراد عليه قد تختلف وتتباين بدرجة كبيرة جدا ، وهذه الاستجابات يقوم المختبر بملاحظتها ومقارنتها وتفسيرها .

وقد وضعت قواعد لتصحيح اختبار بندر إلا أنه عند استخدام الاختبار في أغراض تشخيصية لابد أن يعمل مصححه ومفسره على التكامل النوعي لنتائج الاختبار مع أدوات تشخيصية أخرى .

ومن الممكن معالجة درجات الاختبار احصائيا وذلك عن طريق ملاحظة العلامات أو الإشارات التي تميز جماعة محك معينة ، فمثلا من الممكن تحديد الاستجابات التي تميز العصائيين من العاديين والتي منها مثلا : تكديس الأشكال في نصف الصفحة فقط ، أو العد بصوت مرتفع عند نقل الأشكال ،

وهكذا . وقد اثبتت دراسات أجريت على هذا الاختبار أن هذه الاشارات وغيرها كانت مميزة فعلا بين العاديين والعصابيين (٩) .

٢ - اختبار بقع الحبر لرورشاخ Rorschach Inkblots

وهو أكثر الاختبارات الادراكية غير محددة البنية أو الاسقاطية أهمية واستعمالا وتعرضا للمناقشة والجدل . وقد وضعه هرمان رورشاخ Hermann Rorschach وهو سويسرى الجنسية ، فى عام ١٩٢١ . وعلى الرغم من أن بقع الحبر كانت تستعمل من قبل فى دراسة التخيل وبعض الوظائف العقلية الأخرى فقد كان رورشاخ أول من استعملها فى البحث التشخيصى للشخصية ككل .

ويتكون الاختبار من عشرة اشكال من بقع الحبر ، خمسة منها تتضمن درجات من اللون الرمادى والأسود فقط واثنان باللون الأسود والأحمر ، أما الثلاثة الباقية فهى متعددة الألوان ، ولسهولة استخدام الاختبار وضعت الصور على بطاقات أبعادها ٧ × ٩ ١/٢ بوصة .

وعند تطبيق الاختبار تقدم للمفحوص بطاقة ثم يسأل أن يتكلم عما يراه ، أى عما يمكن أن تمثله بقعة الحبر ويقوم المختبر ، الى جانب تسجيل استجابات المفحوص لكل بطاقة بملاحظة زمن الاستجابة ، والوضع الذى يمسك به المفحوص البطاقة ، وملاحظات المفحوص التلقائية وتعبيراته الانفعالية ، كما يلاحظ أى سلوك عرضى يصدر عن المفحوص أثناء جلسة الاختبار . وبعد تقديم البطاقات العشر للمفحوص يسأل المختبر المفحوص بطريقة منظمة عن الأجزاء والجوانب المختلفة للبقعة تلك الأجزاء والجوانب التى أعطى لها المفحوص ارتباطات . وكذلك يعطى المفحوص فرصة ليوضح ويشرح استجاباته السابقة . ويعكس الترتيب الذى تقدم به البطاقات رغبة رورشاخ نفسه فى ادخال نظام نفسى يضمن بقاء استثارة المفحوص على مستوى ممكن .

ومنذ وضع رورشاخ اختباره . أدخلت عليه تغييرات طفيفة عن طريق باحثين آخرين ، وحدث تحول فى الاهتمام ، فبعد أن كان التركيز على التصنيف الى أنواع امراض نفسية مختلفة ، أصبح التركيز على وصف العوامل الدينامية فى الشخصية Psychodynamics ، وقد أحدث بيك

(٢٢ ، ٢٢) Beck . فى الولايات المتحدة ، تغييرات بسيطة فى تصحيح الاختبار ، ووضع له معايير أولية . الا أن كلوبفر Klopfer أحدث تغييرات أساسية فى نظام تصحيح الاختبار . وسى مصر قام سيد غنيم بدراسات استخدم فيها الرورشاخ على عينات ، كما أجرى دراسة سابقة عن مدى صلاحية الرورشاخ لقياس الذكاء (٧) ووضع مع هدى براده معايير مصرية للرورشاخ (٧ ، ٨) . وكذلك أجرى فاروق أبو عوف دراسة مقارنة عن مدى الافادة من الرورشاخ فى التمييز بين حالات العصاب القهرى والأسوياء (١٠) ، ودراسة أخرى عن مدى صلاحية الرورشاخ لقياس الابتكارية (١١) ، ودراسة ثالثة استخدم فيها المقياس فى التمييز بين القادة وغيرهم (١٢) .

ويعتمد تصحيح الاختبار على قواعد لها طابع موضوعى ، وهناك عدد من الدرجات الرئيسية التى يمكن أن توضع ضمن الأقسام الرئيسية الثلاثة التالية التحديد المكاني ، والمحددات ، والمحتوى . وتدل درجات التحديد المكاني عما اذا كانت الاستجابة تستعمل البقعة كلها (W) ، او الى قسم فرعى يشيع ادراكه (D) ، او تفاصيل غير مألوفة او ينذر الاشارة اليها (Dd) . اما المحددات فهى الشكل . او اللون ، او التظليل التى أخذها المفحوص فى اعتباره . فمثلا تعطى الاستجابة عن الحركة (M) عندما يصف المفحوص اناسا فى حركة و (CF) عندما تعتمد الاستجابة على الشكل واللون ، مع اعتبار اللون أكثر أهمية فى تحديد الاستجابة . ويلاحظ المصحح أيضا مدى ملاءمة الاستجابة لشكل بقعة الحبر ، فإذا كانت الأشكال تدرك بطريقة مألوفة تعطى (F) ، وإذا كانت دقيقة بدرجة غير عادية (F+) . اما اذا كانت بعيدة الصلة ببقعة الحبر او غير محددة ولا مقبولة فتعطى (F-).

اما درجة المحتوى فتسجل ما اذا كانت الاستجابة تشير الى اشخاص (H) واجزاء من اشخاص (Hd) ، او ملابس (G) وهكذا .

وتوجد درجة لدى الشيوع (P) وتعطى على أساس تكرار و شيوع استجابات مختلفة بين الناس ، أما درجة الأصالة (Q) فتشير الى الاستجابات النادرة ولكن تعطى تفسيراً مقبولا لبقع الحبر .

ويقوم تحليل استجابات الرورشاخ بعد ذلك على أساس العدد النسبى

الاستجابات الواقعة فى الأقسام المختلفة . وكذلك على النسب والعلاقات فيما بين هذه الأقسام المختلفة . ومن أمثلة التفسيرات الكيفية التى شاع استعمالها بالنسبة الى درجات الرورشاخ ربط استجابات « الكل » بمستوى التفكير ، وربط استجابات « اللون » بالانفعالية والحركة الانسانية بالتخيل وحياة الوهم . وفى الاستعمال العساذى للرورشاخ يعطى تأكيد رئيسى على الوصف الكلى للشخص ، والذى ينظم الاكلىنىكى فيه الناتج فى صورة متكاملة أخذا فى الاعتبار علاقات الدرجات المختلفة بعضها ببعض الآخر ، كما يستفاد فى الواقع من أية معلومات يحصل عليها المختبر عن المفحوص ، سواء كانت من اختبارات أخرى أو من المقابلات أو من سجلات تاريخ الحالة ، يستفاد من هذا كله فى اعداد ذلك الوصف الكامل للشخص .

وفى الامكان وضع معايير لدرجات الرورشاخ ، الا انها يجب أن تقوم على جماعات فعلية ، وعلى الرغم من أن الرورشاخ قابل للتطبيق ابتداء من سنوات ما قبل المدرسة الى سن الرشد ، فان معايير الأساسية كانت مشتقة اساسا من جماعات من الكبار . وقد بذلت جهود حديثة لتوسع مجالات استعمال الاختبار فى مستويات اعمار متباينة (١٨ : ٥٧١) .

وقد تعرضت بعض الافتراضات التى يقوم عليها تصحيح الرورشاخ لكثير من النقد والشك نتيجة عدد متزايد من نتائج الأبحاث . فمثلا أثبتت دراسة قورنت فيها بطاقات الرورشاخ فى صورتها الملونة بالبطاقات ذاتها معروضة بحيث تلقى فيها الألوان ، أثبتت هذه الدراسة أن اللون ذاته ليس له ذلك التأثير المنسوب اليه على خصائص الاستجابات على بطاقات الرورشاخ (١٨) كما اشارت دراسات أخرى الى أن الاستعداد اللغوى يؤثر فى عدد من درجات الرورشاخ كانت تفسر عادة على انها مؤثرات على سمات معينة للشخصية ، وكذلك وجدت بعض الدراسات ارتباطا عاليا بين درجة تعقد استجابات الرورشاخ وبين درجات المفحوصين على اختبار استعداد لغوى ، وبين تعقد هذه الاستجابات والسن والمستوى التعليمى (١٨) .

ومن العوامل التى تزيد من تعقيد تفسير درجات الرورشاخ مجموع عدد الاستجابات ، والذى يعرف « بانتاجية الاستجابة » أو (R) فمن الصعوبات التى تمثلها انتاجية الاستجابة مثلا انه يبدو أن هذه الانتاجية - كما هو الحال بالنسبة الى تعقد الاستجابات

كما سبق - متصلة بعوامل مثل السن ، والمستوى العقلى ، والمستوى التعليمى ، بل ان الأكثر من هذا مغزى ان هناك ما يشير الى ان انتاجية الاستجابة ، تختلف اختلافا دالا من ممتحن الى اخر (١٨) . وتوحى هذه النتائج كلها بان انتاجية الاستجابة والتي تعتبر فى ذاتها محددا لكثير من درجات الرورشاخ متأثرة بعوامل خارجية وبعيدة عن العوامل التى يزعم ان الرورشاخ يقيسها .

أما بالنسبة الى صدق الاختبار فان نتائج ما أجرى عليها من دراسات غير مشجعة بصفة عامة على الرغم من كثرتها ، اذ ان معظم هذه الدراسات قد فشلت فى ايجاد أية علاقة دالة بين درجات الرورشاخ وبين مجموعات من الدرجات على اختبارات أخرى او على تقويمات عامة شاملة للشخصية . وقد وجد أن للرورشاخ صدقا تنبؤيا وصدقا مصاحبا ضعيفا عندما درس من حيث علاقته بمحكات مثل التشخيص الاكلينيكي ، وبلاستجابة للعلاج النفسى ، وبالنجاح أو الفشل فى مهن مختلفة تلعب فيها خصائص الشخصية دورا هاما ، ويوجد صراعات ومخاوف وأوهام امكن تمييزها بصورة مستقلة . وحتى الدراسات التى اعطت نتائج ايجابية فقد اتضح ان فيها اخطاء منهجية خطيرة (٢٧ : ٥٦٤ - ٥٦٥) . ومما ترتب على النتائج السالبة لصدق اختبار الرورشاخ ان تغير التركيز فى التصحيح من التصحيح الادراكى التقليدى الى تحليل محتوى الاستجابات . لقد كان رورشاخ ذاته واتباعه يعتمدون على الأسس الادراكية للارتباطات التى يعطيها الشخص لبقع الحبر ، وكانت الافتراضات الكامنة وراء هذا الاتجاه هى ان استجابات المفحوص لبطاقات الرورشاخ تدل على استجاباته الادراكية العادية ، وأن سمات الشخصية تؤثر فى الادراك ، وهذا بالنسبة الى التصحيح الادراكى التقليدى ، أما فى تحليل المحتوى فان التركيز يكون على ما يدركه المفحوص فى بقع الحبر ، ومن الممكن ان ننظر الى هذا المحتوى ونعامله كما ننظر الى أى معلومات او مادة يقدمها المفحوص فى جلسة اكلينيكية . ويبدو أن النتائج فى اتجاه تحليل المحتوى واحدة وعشرة أكثر من النتائج المعتمدة على طريق التصحيح الادراكى .

فى ضوء هذه الدراسات والنتائج المتاحة حتى الآن فانه يبدو أن اختبار الرورشاخ يجب أن يستعمل على أساس انه أداة مساعدة فى المقابلة

يستعين به الاكلينكى ذو الكفاءة العالية والدربة الكافية على أعظانه وتصحيحه وتفسيره . والواقع ان الأساس الذى يقف عليه المدافعون عن استعمال الرورشاخ - كما يقول كرونياك - (٢٧ : ٢٦٤) - هو النجاح الذى لقيه الاختبار فى حالات فربية .

هذا ، ومن الطريف الذى له مغزى بعيد ان تشير بايجسار الى ان عوامل التفاعل الاجتماعى بين المفحوص والمختبر تؤثر فى استجابات المفحوص على الرورشاخ ، فان قدرا لا بأس به من نتائج البحوث يوضح ان التفاعل الاجتماعى بين مفحوص معين ومختبر معين يؤثر على جوانب كثيرة من استجابات المفحوص من جهة وعلى تفسير الممتحن لهذه الاستجابات من جهة أخرى (٤٥ ، ٤٦ ، ٥٧) ، وهذا مجال خصب وفسيح للدراسات العربية تكمل وتنمى ما بدأه الباحثون فى الغرب من عمل جاد حول الرورشاخ .

ونحب ان نختم هذا العرض الموجز لاختبار الرورشاخ بنظرة الى قدرة الاختبار على تمييز السواء او جوانب القوة الشخصية ، ذلك لأن هناك قدرا كبيرا من نتائج البحوث يشير الى ان اوصاف الشخصية المعتمدة على سجلات الرورشاخ منحازة الى اتجاه عدم التوافق او عدم السواء . ويرجع والين (٦٠ : ٢١٤) Wallen هذا الى ان طريقة الرورشاخ نشأت وتطورت

فى معظمها متصلة بالأشخاص غير الأسوياء ، والاحتمالات الراجعة ، فى رأى والين ، ان هذه الطريقة ليست مناسبة لتحديد جوانب القوة والسواء فى الشخصية . ويستدل والين على هذا بالدراسة المستفيضة التى قامت بها رو (٥٢) Roe والتى استعملت فيها الرورشاخ لدراسة بعض العلماء

البارزين ، والتى قررت فيها عجبها من ان سجلات الرورشاخ التى جمعناها عن هؤلاء العلماء لم تعطها أية اشارة الى الحياة الواقعية او الأداء الواقعى لهؤلاء العلماء المفحوصين . وهى تشير الى ان هذه السجلات توضح فى الحقيقة وجود مشكلات جادة فى حياة هؤلاء المفحوصين ولكنها ، اى السجلات لا تعطى ما يوضح ان هؤلاء المفحوصين استطاعوا ان يواجهوا هذه المشكلات بنجاح وفاعلية . وان اتجاه استعمال الرورشاخ فى دراسة هذه الجوانب السوية والقوية فى الشخصية يمثل أيضا مجالا جديرا باهتمام الباحثين العرب .

٢ - اختبار يقع الحبر لهولتزمان

The Holtzman Inkblot Technique

نصم هذا الاختبار على نمط الروشاش ، الا ان هولتزمان (٢٧) حاول فيه ان يتجنب العيوب الفنية الرئيسية للروشاش : والتغيرات التي ادخلها هولتزمان على الاختبار سواء في مواد المثيرات أو الطريقة تغيرات كبيرة وشاملة تجعل من الممكن ان ندرسه ونقومه دون الاشارة الى الخصائص النظرية أو العملية للروشاش . ويتضمن اختبار هولتزمان مجموعتين متماثلتين من سلاسل البطاقات ، تضم كل مجموعة عدد ٤٥ بطاقة . والمطلوب من المفحوص ان يعطى استجابة واحدة لكل بطاقة وعلى هذا ثبت انتاجية الاستجابة بالنسبة الى كل مفحوص وتتلافى كثيرا من عيوب التصحيح التقليدي للروشاش وطريقة اجراء الاختبار وتصحيحه مقننة تقنيا جيدا ومشروحة بشكل واضح . وللاختبار معايير لعينات سرية تبدأ اعمارها من الخامسة الى الرشد ، وعينات غير سرية ، ويبدو ثبات المصحح مقبولا ، ونتائج الثبات ، سواء الاتساق الداخلي ، أو التكافؤ مشجعة ، وان كانت دراسات الصدق ، ما زالت قليلة . وللاختبار صورة جماعية تعطى درجات قريبة من الدرجات التي يعطيها الاختبار عندما يطبق بصورة فردية .

٤ - اختبار تفهم الموضوع

Thematic Apperception Test (TAT)

وهناك اختبار آخر من الاختبارات الإدراكية غير محددة البنية ، أو الإسقاطية ، الا انه يختلف عن الاختبارات التي سبقتنا مناقشتها في ان مثيراته تعتبر أكثر تحديدا في البنية نسبيا ، ولأنه يتطلب استجابات وأنشطة أكثر تعقيدا وأكثر خضوعا للتحكم العقلي من مثيلاته من الاختبارات السابقة . ويعتمد تفسير هذه الاستجابات عادة على تحليل محتوى ذي طابع كيفي وذلك تمشيا مع الاتجاه الغالب الى انه من الأكثر فائدة ان نفس الاختبارات الإسقاطية عامة على أساس تحليل المحتوى ، ويعطى اختبار تفهم الموضوع وغيره من الاختبارات التي تعتمد على القصص مجالا أكثر اتساعا لتحليل المحتوى من اختبار الروشاش .

وكان اول ما نشر عن اختبار تفهم الموضوع مقال نشره هنرى موري H. Murray وزميلته Morgan فى عام ١٩٢٥ ، واستمرت الأبحاث التى قام بها موري وزملاؤه فى العيادة النفسية فى جامعة هارفارد بالولايات المتحدة ثم ظهرت فى كتاب فى عام ١٩٢٨ (٥٠) وقد استعمل هذا الاختبار منذ ظهوره استعمالا واسعا فى الممارسة الاكلينيكية وفى البحث ، كما يعتبر نموذجا اقتدى به كثير من اختبارات الشخصية .

تتكون مواد الاختبار من ١٩ بطاقة تحتوى كل منها صورة غامضة باللونين الأبيض والأسود ، بالاضافة الى بطاقة خالية من أى صورة . ويطلب من المفحوص ان يكون قصة تناسب كل صورة ، ذاكرة ما أدى الى ما تمثله الصورة ، واصفا ما يحدث وما تحس به وتفكر فيه الشخصيات الممثلة فى الصورة ، وعلى أن يذكر النتيجة التى يتوقع أن ينتهى اليها ما يصفه . اما بالنسبة الى البطاقة البيضاء التى ليس عليها أى صورة ، فان التعليمات تطالبه بان يتخيل صورة ما على البطاقة ويصفها ثم يذكر قصة عنها . وتتطلب الطريقة الأصلية كما حددها موري جلستين لكل منهما ساعة ، مع استعمال ١٠ بطاقات فى كل جلسة ، وتختار البطاقات المحجوزة للجلسة الثانية على أساس انها غير مألوفة وشاذة وذات طابع درامى ، والتعليمات المصاحبة هى ان يطلق المفحوص لخياله العنان وتوجد اربعة مجموعات من الاختبار ، وهى متداخلة ، وتتكون كل مجموعة من ٢٠ بطاقة ، منها مجموعة للاولاد ، واخرى للبنات ، وثالثة للذكور فوق سن ١٤ ، ورابعة للاناث فوق سن ١٤ ، ويستعمل معظم الاكلينيكين مجموعات مختصرة من بطاقات منتقاة ، وقادرا ما يعطى اكثر من عشر بطاقات لمفحوص واحد .

وعلى الرغم من ان الاختبار اصلا هو اختبار فردى شغوى يعطى فى الموقف الاكلينيكى ، فمن الممكن اعطاؤه كتابة ، أى يقوم المفحوص فيه بالكتابة بدل الكلام كما يمكن اعطاؤه فى جماعة : وتشير نتائج بعض البحوث الى انه فى حالة اعطائه جماعيا وكتابة يكون انتاج المواد ذات المعنى والمغزى اكثر سهولة ويسرا (١٨ : ٥٠١) وهناك ميل الى تأكيد اهمية العوامل والمؤثرات الاجتماعية فى الموقف الاختبارى بالنسبة الى اختبار تفهم الموضوع . كما لاحظنا اثر هذه العوامل والمؤثرات ونوعنا بها بالنسبة الى اختبار

الرورشاخ ، ونكرر ان هذا ايضا مجال طيب لمزيد من البحث النفسى الاجتماعى .

اما فى تفسير قصص اختبار تفهم الموضوع فيحدد المختبر اولا البطل hero وهو عبارة عن الشخصية ، ذكرا كانت ام انثى ، التى يتوحد معها المفحوص ثم يحلل محتوى القصص اساسا وفق ما حدده من حاجات needs . وضغط Press . وقد سبقت الاشارة الى الحاجات عند مورى عند الكلام عن مقياس التفضيل الشخصى الذى وضعه ادواردز ، ومن امثلة هذه الحاجات الانجاز achievement والاعتداء aggression ، والاعالة nurturance ، والجنس . الا أن مورى يعطى حاجات معينة اهمية كبيرة فى التفسير مثل : تجنب الأذى، البحث عن الكمال، الاعتداء، السيطرة، العدوان الموجه نحو الذات ، الاستنجاذ ، السلبية ، الجنس (٩ : ١٤٤ وما بعدها) .

ومع معرفة البطل الرئيس ، ودراسة الحاجات والدوافع الأساسية للسلوك لابد من معرفة ما اسماء مورى الضغط أو الضغوط ، وهى تشير الى القوى البيئية التى قد تيسر أو تعوق اشباع الحاجات ، ومن امثلة هذه الضغوط ما يشار اليه من التعرض للهجوم أو النقد بواسطة شخص آخر ، وتقبل العطف والمحبة ، والتعرض للخطر المادى أو الجروح المادية .

وعند تقدير أهمية أو قوة حاجة أو ضغط معين عند الشخص يعطى اهتمام خاص بعمق وحدة وتكرار حدوث هذه الحاجة أو الضغط فى القصص المختلفة بالإضافة الى الاهتمام بارتباط الحاجة أو الضغط بصورة معينة بصفة خاصة . والاقتراض القائم هو أن أية استجابات مخالفة للاستجابات الشائعة لكل صورة يغلب أن يكون لها - أى الاستجابات المخالفة - مغزى بالنسبة الى المفحوص .

وعلى الرغم من أن تفسير اختبار تفهم الموضوع كان اساسا تفسيريا كيفيا وانطباعيا كما سبقت الاشارة ، فان كروبناك (٢٧ : ٥٧١) يذهب الى أنه من الممكن انشاء وتطوير نظم موضوعية لتفسير هذا الاختبار ، وفى رأيه ان هناك عددا كبيرا من المتغيرات العامة المشتركة التى يمكن ان نلاحظ قوتها

تقريبا في كل اداء على الاختبار ، ومن امثلتها : ادراك السلطة ، ورد الفعل
للاعمال الصعبة جدا ، والاصالة : والاعتماد على الحظ وتأثير السحر .
وتوجد امثلة متعددة لمحاولات التصحيح الكمية التي اعطت ثباتا عاليا
للمصححين ، الا ان تطبيق هذه الأساليب يستغرق وقتا طويلا لهذا ندر
استعمالها في المجال الاكلينيكي .

وقد نشر قدر لا بأس به من المعلومات المعيارية عن الاختبار فيما يتصل
بأكثر الاستجابات المميزة لكل بطاقة شيوعا ، وبالطريقة التي تدرك بها كل
بطاقة ، والموضوعات المذكورة ، والأدوات التي تنسب الى الشخصيات ،
وسرعة الاستجابات ، وطول القصص ، وعلى الرغم من ان هذه البيانات
الخاصة بالمعايير تمثل اطارا عاما لتفسير الاستجابات المفردة ، فان معظم
الاكلينيكيين يعتمدون اعتمادا رئيسيا على « معايير ذاتية » بنوها من خبرتهم
الشخصية في اعطاء الاختبار .

وقد استعمل اختبار تفهم الموضوع استعمالا واسعا في بحوث
الشخصية ، ومن امثلة استعماله في مصر بحث احمد عبد العزيز سلامة (٢) .
وركز بعض هذه البحوث على الافتراضات التي تقوم عليها تفسيرات الاختبار
مثل توحد الذات مع البطل ، والمغزى الذاتي للاستجابات غير المألوفة .
وعلى الرغم من ان هذه البحوث لا تعطى تأييدا للصدق التنبؤي والصدق
التلازمي ، فانها تضيف الى صدق التكوين الفرضي لتفسيرات اختبار تفهم
الموضوع . ومن الافتراضات الأساسية الأخرى التي يشارك فيها هذا الاختبار
الاختبارات الاسقاطية الأخرى ان ظروف المفحوص العاطفية والدافعية وقت
الاختبار تؤثر على استجابات المفحوص لموقف الاختبار غير محدد البنية .
ويتوافق قدر كبير من نتائج الدراسات التجريبية يشير الى ان ظروفًا مثل
الجوع ، والحرم من النوم ، والحبوط الاجتماعي . ومعاناة الفشل في
موقف اختياري سابق ، كلها تؤثر على استجابات اختبار تفهم الموضوع .
وبينما يؤكد هذا صحة الفرض الاسقاطي فان حساسية هذا الاختبار لهذه
العوامل المؤقتة يزيد من صعوبة الكشف عن سمات الشخصية الثابتة
(١٨ : ٥٠١) .

ولمنا نلاحظ ان اختبار تفهم الموضوع قد صمم أساسا بحيث

يشمل جوانب الشخصية جميعا ، كل الأفكار والمشاعر والسلوك ، وعلى هذا فلا يستطيع أن يسير موضوعا أو جانبا واحدا بعمق ، ولهذا فقد انشئت اختبارات تفهم موضوع تركّز على جانب بعينه ، أى أنها تصمم لاستدعاء استجابات تدور كلها حول مسألة واحدة وقد بدأها مورى وليكرت Murray and Likert فى دراسة بين العمال والادارة وذلك عن طريق تقديم صورة تمثل تظاهرات مع رجال الأمن .

أما المحاولة المستفيضة المتعمقة للاستفادة من هذه الاختبارات المركزة فى تفهم الموضوع فكانت محاولة ماكيلاند ورفاقه (٤٧ ، ٤٨) وكذلك محاولة اتكنسن (١٩ ، ٢٠) .

لقد اختار ماكيلاند أربع صور ، منها اثنتان من اختبار تفهم الموضوع ، كان القصد استدعاء الاتجاهات نحو الانجاز . وكانت هذه البطاقات الأربع تمثل موقف عمل (رجل امام ماكينة) ، وموقف مذاكرة (تلميذ يجلس الى مكتبه ومعه كتاب) ، وصورة أب وابن ، وصورة ولد يبدو مستغرقا فى أحلام اليقظة . ويقاس دافع الانجاز فى كل قصة تعبر عن التنافس والكذاح للوصول الى مستوى أو تحقيق هدف أما اتكنسن فقد استعمل اختبار آخر هو French Test of Insights وليذات الغرض ، أى قياس الدافس الى الانجاز ، ويقوم الاختبار على إعطاء وصف موجز للسلوك مثل : « يعطى الفتى بل فرصته لزميله ليفوز » ، ويسأل المفحوص أن يعطى تفسيرا لهذا . ويتضمن الاختبار عددا من مثل هذه العبارة وتتكرر الدرجة عليه من عدد المرات التى تذكر فيها الرغبة فى الانجاز باعتبارها دافعا . وتمثل هذه الدراسات خطأ من الأبحاث الغنية التى نبعت أصلا من اختبار تفهم الموضوع .

٥ - اختبار تفهم الموضوع للأطفال :

Children's Apperception Test (CAT)

وعلى الرغم من أن اختبار تفهم الموضوع الأصلي وضع بحيث يصلح فى التطبيق على أطفال يبلغون من العمر أربع سنوات فإن اختبار تفهم الموضوع

للأطفال قد صمم خاصة ليطبق على الأطفال ممن تتراوح أعمارهم بين ثلاث وعشر سنوات . ويستعمل هذا الاختبار على بطاقاته صور الحيوانات بدل البشر على أساس افتراض أن الأطفال يسقطون على الحيوانات بسهولة أكثر من إسقاطهم على البشر ، إلا أن صور الحيوانات موضوعة في مواقف بشرية ، وهذه الصور مصححة بحيث تستدعي عند الأطفال أوهاما وتخيلات تتصل بمشكلات التغذية والأنشطة الفنية الأخرى ، ومشكلات العلاقات بين الأشقاء ، والعلاقات مع الأبوين ، والعدوان ، وعمليات الإخراج وضبطها ، والخبرات الأخرى الخاصة بالطفولة ، وعلى العكس من افتراض واضعى هذا الاختبار ، أثبتت دراسات متعددة أن استعمال الصور البشرية وصور الحيوانات لم يعط اختلافات تذكر ، بل ظهر غالباً أنه في حالة استعمال الصور البشرية كانت انتاجية الأطفال من المواد ذات المغزى الاكلنيكى اكبر منها في حالة استعمال صور الحيوانات ، وفي ضوء هذه النتائج صمم واضعو الاختبار صورة بشرية منه ، هي CAT-H ، للاستعمال مع الأطفال الكبار ، وخاصة من يزيد عمرهم العقلى عن عشر سنوات ، وذهبوا الى أن أيا من الصورتين ، البشرية والحيوانية - يكون أكثر فاعلية حسب سن الطفل وخصائص شخصيته .

(ج) الاختبارات اللفظية غير محددة البنية :

على الرغم من الاختبارات غير محددة البنية التى نوقشت حتى الآن تتطلب استجابات لفظية بينما المثيرات غير لفظية ، فإن هناك نوعاً من الاختبارات غير محددة البنية أو الاسقاطية تعتمد على مثيرات لفظية كما تتطلب استجابات لفظية سواء كانت منطوقة أو مكتوبة ، وهى مناسبة للتطبيق الجماعى بالكتابة . وتذكر فيما يلى مثالين من هذه الاختبارات .

١ - اختبار تكلمة الجمل :

وتستعمل اختبارات تكلمة الجمل فى اختبارات الاستعداد اللغوى إلا أنها اختبارات تكلمة الجمل ذات البنية غير المحددة تختلف عن الأولى فى أنها تتطلب استجابات وتكملات متنوعة ومتباينة على أساس أنها شروع من

الاختبارات الاسقاطية • وغالباً ما تعطى الكلمة الأولى ، أو بدايات الجمل ، وعلى المفحوص أن يكتب النهاية ، ومن أمثلتها ما يلي :

أنا أحس ...

ما يضايقني ...

النساء ...

وتصاغ بدايات الجمل بحيث تستدعي استجابات تقتصر بمجال الشخصية المراد التعرف عليه ، وتعتبر هذه المرونة التي يتمتع بها أسلوب تكملة الجمل بميزة هامة تركيزه للاستعمال الاكلينيكي وفي البحث • وقد نشرت صور مقننة من هذه الطريقة للاستعمال في مجال أوسع ، منها قائمة روتر للجمل الناقصة Rotter Incomplete Sentence Blank

وتتكون من عدد ٤٠ بداية للجمل ، وتنص التعليمات الموجهة الى المفحوص على أن يكمل هذه الجمل لينبهر عن مشاعره الحقيقية ، والا يترك أية جملة ويتأكد من أنه يكون جملة كاملة ، ومن أمثلة هذه الجمل الناقصة •

أنا أحب ..

ان خوفي الأكبر ..

أحياناً ...

أنا أرغب ...

والذي ...

وتقدر كل تكملة على مقياس ذي سبع نقاط حسب درجة التوافق أو عدم التوافق التي تدل عليها التكملة • ويعطى دليل الاختبار عينة من التكملات والدرجة المناسبة لها مما يسمح بقدر مناسب من التقدير الموضوعي للدرجة، ويمثل مجموع التقديرات درجة كلية للتوافق يمكن استعمالها في أغراض الانتقاء • ومن الممكن أن يفحص محتوى الاستجابة للتعرف على ما قد يكون فيها من مؤشرات ذات دلالة تشخيصية ، وتشير دراسات صدق الاختبار الى امكانياته الواسعة في الاستعمال •

النفس التجريبي والقياس العقلي المبكرة للتعرف على عمليات التفكير ، ثم اتسع مجاله في الميدان الاكلينيكي عندما انتشر أسلوب التحليل النفسي ولعل من أهم الاختبارات المعتمدة على تداعى الكلمات والتي تزايدت أهميتها منذ وضع اختيار كنت - روزانوف للتداعى الحر Kent — Rosanoff Free Association Test والذي وضع فى عام ١٩١٠ ، واعتمد تماما على التصحيح الموضوعى وعلى المعايير الاحصائية . وتتكون الكلمات المثيرة من مائة كلمة محايدة شائعة اختيرت لأنها تميل الى ان تدعاه مثيرات متشابهة عند الناس بصفة عامة . مثلا يميل معظم الناس الى ان يستجيبوا لكلمة طاولة بكلمة كرسى ، ولكلمة ظلام بكلمة نور ، وأعدت مجموعة من جداول التكرار - واحد لكل كلمة مثير - مبينة عدد المرات التى أعطيت فيها الاستجابة لعينة تقنين عددها الف راشد سوى .

وفى تصحيح اختبار كنت - روزانوف ، تستعمل قيمة وسيط التكرار للاستجابات التى أعطاها المفحوص للمائة كلمة المثيرة على أنها « مؤشر للشروع » ، وتعتبر الاستجابات التى لا توجد فى جداول المعايير « فردا » ، وقد بينت مقارنات الذهانين بالأسوياء ان الذهانين أعطوا عدد استجابات « فرد » أكثر ، كما حصلوا على « مؤشر للشروع » منخفض عن الأسوياء . الا ان القيمة التشخيصية للاختبار بدأت تتناقص مع التحقق التدريجى من ان تكرار الاستجابات يختلف باختلاف السن ، والمستوى الاجتماعى والاقتصادى والمستوى التعليمى وغيرها من العوامل . وعلى هذا فان التفسير السليم للدرجات على الاختبار يتطلب معايير الجماعات الفرعية ، وكذلك معلومات اضافية عن المفحوص . وعلى أية حال فان الاهتمام بهذا الاختبار مستمر ومتجدد فى مجالات السلوك اللغوى والشخصية فضلا عن استعماله الاكلينيكية . وقد قام فؤاد أبو حطب بنقل هذا الاختبار الى اللغة العربية وطبقه على عينات مصرية بلغ حجمها ١٠٠٠ مفحوص . راعى له معايير تقبل المقارنة مع المعايير الأجنبية (*) .

(*) فؤاد أبو حطب : المعايير المصرية لاخبار قداى الكلمات : فى بحوث تقنين الاختبارات النفسية ، المجلد الاول (تحرير فؤاد أبو حطب) ، القاهرة ، مكتبة الانجلو المصرية ، ١٩٧٧ .

الفصل السادس عشر

قياس الميل

يرجع الاهتمام بالدراسة النفسية للميل الى الترجية التربوي والمهني، كما يرتبط انشاء الاختبارات التي تقيس هذه الميل بعمليتي الانتقاء والتصنيف المهني، لأن للميل أهمية عملية، ويجب أن تؤخذ في الاعتبار من جانب كل من طالب العمل وصاحب العمل، فالأداء في عمل من الأعمال هو خلاصة للاستعداد والميل معا. وعلى هذا فان قياس هذين المتغيرين، أي الاستعداد والميل، يسمح بدرجة اكبر من الدقة في التنبؤ بالأداء من الاقتصار على واحد منهما فقط. ذلك لأنهما، أي الاستعداد والميل، على الرغم من ارتباطهما فان مستوى مرتفعا في أحدهما ليس يعنى بالضرورة مستوى مرتفعا في الآخر. فقد يكون عند الشخص استعداد للنجاح في نمط معين من النشاط، سواء كان دراسيا، أو مهنيا، أو ترفيهيا، دون أن يكون عنده الميل المناسب لهذا النشاط أو قد يكون عنده ميل لعمل ما ولكنه يفتقد الاستعدادات المطلوبة لهذا العمل (راجع ١ : ٢٧٥ وما بعدها). أما وقد تكلمنا من قبل عن قياس الاستعدادات، فانتا نتناول في هذا الفصل قياس هذا المتغير الآخر وهو الميل.

مقاييس الميل

يعتبر اختبار الميل استفتاء أو استبياناً مطولا يستخدم أسلوب التقرير الذاتي الذي تكلمنا عنه في الفصل الثامن، ويهدف هذا الاستبيان الى الحصول على معلومات بأن نجعل الشخص يصف خصائصه ومميزاته الذاتية وهذا الاختبار أو الاستبيان عبارة عن نوع من المقابلة المكتوبة، ولما كان يتضمن عددا كبيرا من الأسئلة غير المباشرة. فانه يفضل المقابلة الشفوية المباشرة.

والواقع أن طريقة السؤال المباشر للمفحوص لتتعرف على ميله. تعنى أن نسأله عما يميل اليه من أعمال أو دراسات أو ترويح، وهذه الطريقة

تعرضت لنقد شديد في وقت مبكر ، في خلال العقد الثاني من القرن العشرين .
فقد تجمعت دراسات شتى بينت أن الإجابة عن الأسئلة المباشرة فيما يتصل
بالميول كثيرا ما تكون غير ثابتة ، و سطحية ، وغير واقعية (١٨) . وترجع
أسباب هذا إلى أن معظم الناس ليس لديهم المعلومات الكافية عن الأعمال
وموارد الدراسة والأنشطة المختلفة المتضمنة في عمل ما ، ومن ثم فإنهم يكونون
غير قادرين على أن يقرروا ما إذا كانوا سوف يحبون كل ما يشتمل عليه
ما يحبونه من نشاط . وقد يكون ميلهم ، أو عدم ميلهم ، بالنسبة إلى عمل
من الأعمال راجعا إلى قصور ما لديهم من معلومات عما يتضمنه هذا العمل
من تفاصيل واجراءات وتصرفات تجري كل يوم . ومن أسباب عدم الثبات
والسطحية وعدم الواقعية هذه شيوع صور وافكار جامدة ثابتة . عن بعض
الأعمال والمهن ، كما هو الحال بالنسبة إلى المدرس والمهندس والمحامي
والطبيب ، وغيرهم ، هذه الصور والافكار تشيع وخاصة بين الشباب ، عن
طريق وسائل الاعلام المختلفة . فالمشكلة إذن هي أن الفرد ليس في وضع
يسمح له بأن يعرف ميوله في المجالات المختلفة قبل أن ينخرط في سلك العاملين
في هذه المجالات ، وقبل أن تفوت الفرصة للاستفادة من معرفته حين لا يستطيع
تغيير عمله أو تغييره . لهذا كله عدل عن طريقة السؤال المباشر وتحول الاتجاه
إلى الاعتماد على الأسئلة غير المباشرة المتضمنة في اختبارات الميول التي
سبق وصفها بأنها بمثابة المقابلة المكتوبة .

وقد اتبعت في بناء اختبارات الميول وتصحيحها عدة أساليب :

١ - فمنها ما اعتمد في بناء تكوينه على أساس تجريبي empirical
واقعي ، مثل اختبار سترونج للميول المهنية .
Strong Vocational Interest Blank (SVIB)

٢ - ومنها ما اعتمد على التنوع والتعدد ، مثل اختبار كيودر .
Kuder Preference Record

٣ - ومنها أخيرا ما اعتمد على أساس منطقي مثل اختبار لي -
ثورب .
The Lee - Thorpe Inventory

وسوف نتناول كلا من هذه الاختبارات .

أولا : اختبار سترونج للميول المهنية :

نشأ اختبار سترونج مع مجموعة أخرى من اختبارات الميول التي انبثقت عن سيمينار انعقد في العام الدراسي ١٩١٩ - ١٩٢٠ في مؤسسة كارينجي للتكنولوجيا في الولايات المتحدة Carnegie Institute of Technology . حيث كان موضوع هذا السيمينار هو الميول ، الا انه لم يكتب لأي من تلك الاختبارات الاستمرارية والازدهار مثل ما كتب لسترونج الذي تعرض منذ نشأته للمراجعة والاتساع والبحث . وقد قام بترجمته الى اللغة العربية عطية محمود هنا .

وتتكون الصورة الحالية لاختبار سترونج ، في أصله الاجنبى ، والمنشورة في عام ١٩٦٦ من عدد ٢٩٩ عنصرا مجمعة في ثمانية اجزاء . وفي الاجزاء الخمسة الاولى يسجل المفحوص تفضيله بأن يضع دائرة حول احد الحروف L (يحب Like) او I (محايد Indiffirent) او D (لا يحب Dislike) بالترتيب . وكل واحد من هذه الاجزاء الخمسة مختص بواحد من الأقسام التالية : المهن ، المواد الدراسية ، الأنشطة المسلية ، وأنشطة متنوعة (مثل اصلاح الساعات . او التعبير عن الآراء والأحكام علنا بصرف النظر عن نقد الآخرين) . والخصائص الغريبة وغير المألوفة للناس . اما الاجزاء الثلاثة الأخرى من الاختبار فتتطلب من المفحوص ان يرتب أنشطة معينة حسب تفضيله ، ويقارن ميله في ازواج من العناصر ، وأن يقدر قدراته الحالية وخصائصه ومميزاته .

يصحح الاختبار بمفتاح خاص لكل مهنة ، وحاليا يوجد ٥٤ مقياسا مهنيا لتصحيح الصورة الخاصة بالرجال من الاختبار و ٢٢ مقياسا مهنيا لتصحيح الصورة الخاصة بالنساء ، وتنشأ مقاييس جديدة من وقت لآخر كلما تجمعت معلومات وبيانات عن جماعات عمل أو مهن أخرى وعند انشاء هذه المقاييس المهنية قورنت استجابات أشخاص ناجحين في مهنة باستجابات الرجال بصفة عامة أو النساء بصفة عامة . وتتكون جماعة المرجع هذه على أساس افتراض هو أن معظم مقاييس

المهن فى اختبار سترونج تتصل بالمهن والوظائف النجارية العليا . الا انه قد ثبت ان استخدام جماعة مرجع ممثلة بكل مجتمع الذكور كان اقل نجاحا فى التمييز بين المهن المفردة . فميل المهنيين ورجال الأعمال كجماعة تختلف اختلافا شاسعا عن ميل العمال المهرة ، فعندما استعملت جماعة المرجع العامة ادى هذا الى طمس وعدم وضوح الفروق بين المهن ذات المستوى المرتفع .

وقد اختبرت عناصر الاختبار لكل مقياس مهني وأعطيت لها أوزان على أساس اختلاف تكرار اختيار الرجال فى المهنة المتصلة بالمقياس والرجال بصفة عامة . فمثلا فى مقياس المحامى يدل وزن + ١ على ان الاستجابة تعطى اكثر فيما بين المحامين عنها بين الرجال عموما ، بينما يدل وزن - ١ على ان الاستجابة تعطى اقل فيما بين المحامين عنها بين الرجال عموما ، والاستجابات التى لا تميز بين المحامين والرجال بصفة عامة ، لا تظهر فى مقياس المحامى . بصرف النظر عن مقدار اختيار المحامين لها . وتتكون الدرجة الخام للمفحوص على ان كل مقياس مهني من المجموع الجبرى لأوزانه الموجبة والسالبة وتحول هذه الدرجات الخام الى درجات معيارية متوسطها ٥٠ وانحرافها المعيارى ١٠ حسب توزيع الدرجات فى كل جماعة محك مهنية occupational criterion group وجماعات المحك المهنية هذه كانت تتكون عادة من ٢٠٠ شخصا من الرجال تتراوح اعمارهم بين ٢٥ و ٥٥ ، ملتحقين بأعمالهم لمدة ثلاث سنوات او اكثر ، وقد قرروا انهم راضون عن أعمالهم .

وقد أعطيت تقديرات حروف لزيادة سهولة وتفسير الدرجات المعيارية، فمثلا التقدير (١) (A) يمثل درجة عند او اكبر من - ١/٢ ع (انحراف معيارى)، أى نقطة فوقها نحو ٦٩٪ من جماعة المهنة المعينة قد أعطوا استجاباتهم . وبالمثل فان الحد الأدنى لتقديرات ب (B) قد وضع عند - ٢ ع ، وتقدر الدرجات اقل من - ٢ ع بالتقدير حـ (C) . وعلى هذا فان التقدير حـ يعنى ان الشخص قد حصل على درجة على هذا المقياس بالذات اقل من ٩٨٪ من العينة المماثلة . وعلى أى حال فانه جدير بالملاحظة أنه مهما كانت الصورة التى تمثل بها الدرجات على اختبار سترونج ، فان هذه الدرجات معتمدة على الاختلافات بين ميل جماعة مهنية وميل الرجال عموما . وعلى هذا فان درجة مرتفعة على مقياس الطبيب مثلا تدل على ان ميل الشخص يختلف

عن ميول الرجال بصفة عامة في نفس الاتجاه الذي تختلف فيه ميول الأطباء .

هذا هو المقصود من قولنا ان سترونج قد بنى اختبارها على أساس تجريبي على صرف (٢٧ : ٤٠٦) ، دون ان يعتمد الا على قدر ضئيل جدا من المسلمات والافتراضات النفسية ، واقام تصحيحه للاختبار - كما لاحظنا - على أساس الارتباطات الملاحظة للاستجابات مع المحك .

ومن الممكن ان يصحح اختبار سترونج بالنسبة الى الفرد لمهنة واحدة: فمثلا عندما تهدف الى الانتقاء فائنا نحب أن نعرف الى أى مدى تتشابه ميول طالب الوظيفة مع أولئك الناجدين من العاملين فيها ، الا ان الاختبار غالبا ما يصحح كله مما يعطى فرصة لدراسة النمط الكلى لميول المفحوص ، ويعطى هذا النوع من التحليل فرصة اوثق للتنبؤ بالميل المهني الغالب .

ويتضمن اختبار سترونج ، الى جانب المقاييس المهنية ، أربعة مقاييس غير مهنية ، هي :

Specialization Level (SL)	مستوى التخصص
Occupational Level (OL)	ومستوى المهنية
Masculinity — Femininity (MF)	والذكورة والأنوثة
Academic Achievement (AACH)	والتحصيل الأكاديمي

وقد انشئ مقياس مستوى التخصص بمقارنة استجابات المتخصصين في الطب باستجابات جماعة من الأطباء عامة . وقد وجد أن هذا المقياس يمكن تطبيقه في مجالات أخرى غير الطب في التعرف على أولئك الذين سوف يجدون متعة في الدراسة المتقدمة التي تتضمن تخصصا دقيقا . أما مقياس مستوى المهنة فيقيس الاختلاف بين ميول العمال غير المهرة في ناحية ورجال الأعمال والمهنيين في الناحية الأخرى . أما مقياس الذكورة والأنوثة فانه يوضح درجة التشابه بين ميول الشخص وميول الرجال والنساء .

أما مقياس التحصيل الأكاديمي . وهو أحدث إضافة الى المقاييس غير

المهنية ، فقد انشئ بمقارنة استجابات جماعات من طلاب المدارس الثانوية الجامعة ممن حصلوا على درجات مرتفعة بجماعات ممن حصلوا على درجات منخفضة • وقد تبين من البحوث التى تلت اضافة هذا المقياس أنه يعكس ميولا للاجتهادات العلمية والعقلية أكثر من ميل للبيع أو الأعمال التجارية (١٨ : ٤٧٠) •

وقد تعرض اختبار سترونج لبرنامج مستمر من البحث الذى اعطى قدرا كبيرا من البيانات المتصلة بثباته وصدقه فقد كانت معاملات الثبات - بالتجزئة النصفية على أساس أرقام العناصر الزوجية والفردية - لكل مقياس فرد ٠.٨٨ ، وكان واحد منها فقط أقل من ٠.٨٠ وكذلك كانت معاملات الاستقرار باعادة الاختبار بعد فترة اسبوعين وكذلك بعد فترة ثلاثين يوما عالية ايضا ، كما أثبتت دراسات طويلة بدرجة استقرار مناسبة لفترة طويلة للدرجات على المقاييس المفردة • وكانت هذه الدراسات تمتد بين ثلاث سنوات وثمان سنوات ، وامتد بعضها الى ٢٠ سنة أو أكثر (٥٨) •

أما صدق اختبار سترونج ، فيوجد قدر كاف من نتائج البحوث يبين أن المفحوصين يميلون الى الالتحاق والاستمرار فى المهن التى حصلوا فيها على درجات مرتفعة على الاختبار •

وبالنسبة الى تلاميذ الثانوى فإنه من الممكن التنبؤ بدرجة حسنة بالمجال أو جماعة المهنة التى سيلتحق بها التلميذ • أما التنبؤ بمهنة يعينها فإنه لا يوثق به تماما الا فى حالة تلك القلة من التلاميذ التى تظهر عندها أنماط محددة ومميزة جدا من الميول •

ويتوافر قدر من البيانات التى توضح وجود ارتباط بين درجات اختبار سترونج والنجاح فى المهنة أو الرضا عنها ، أما الأدلة التى تثبت عكس هذا فهى نادرة ويصعب تفسيرها •

والواقع أن المزايا التى يتمتع بها اختبار سترونج جعلت انستازى تذهب فى تقديره الى القول بأنه • واحد من أنجح الطرق لقياس المتغيرات

غير العقلية ، (١٨ : ٢٦) • وما زال الاختبار خاضعا لمعدد من مشروعات البحث في • مركز البحث في قياس الميول ، •

The Center for Interest Measurements Research

الذى أسس في عام ١٩٦٢ في جامعة مينسوتا بالولايات المتحدة • ويكفى أن نشير إلى أن واحدا من هذه المشروعات يتضمن دراسة تمتد ٢٥ عاما لأكثر من ألف تلميذ عمرهم ١٥ سنة • ومن المتوقع أن تزيد هذه الدراسات من قيمة الاختبار • وقد نقل هذا المقياس إلى اللغة العربية عطية محمود هنا واستخدم في عدد من البحوث المصرية والعربية •

ثانيا : اختبار كيودر للميول المهنية :

Kuder Preference Record-Vocational

يعتبر هذا الاختبار أكثر مجموعة اختبارات كيودر للميول شهرة أو استعمالا • والواقع أن الصور المختلفة ، والطبعات المعادة والمعدلة لاختبارات كيودر للميول تمثل عائلة من الأدوات التي تتخذ لقياس الميول سبلا وزوايا مختلفة ، كما أنها مصممة لأغراض مختلفة ، إلا أن أهمها هو هذا الذي نتكلم عنه • وقد وضع هذا الاختبار كيودر ، واتبع منهجا مختلفا ، بل مضادا ، لمنهج سترونج في اختبار وتصحيح عناصر الاختبار • لقد كان هدفه الرئيسي هو بيان الميل النسبي في عدد صغير من المجالات الواسعة ، وليس في مهن بذاتها ، وقد صيغت عباراته أصلا وجمعت مبدئيا على أساس صدق المحتوى • وقد اتبع هذا تحليل شامل لعناصر الاختبار من تطبيقها على جماعات من تلاميذ الثانوى والكبار • وكان الهدف من تحليل العناصر هذا هو انشاء مجموعات عناصر ذات اتساق داخلي عال وارتباط منخفض بالمجموعات الأخرى وقد تحقق هذا الهدف بشكل جيد في معظم مقاييس الاختبار •

وعناصر اختبار كيودر من نوع الاختيار الإجابى الثلاثى ، وقد تعرفنا عليه في بعض اختبارات الشخصية من قبل - وعلى الخصوص أن يختار بين الأنشطة الثلاثة المذكورة ذلك النشاط الذى يحبه أكثر كما يبين النشاط الذى يلقى عنده أقل تفضيل ، ومن أمثلة هذه الثلاثيات ما يلى ، وهى مختارة من

الصورة العربية للاختبار التي أعدها الدكتور أحمد زكى صالح (ص ٧ من كراسة الأسئلة) .

١٦ - تعمل نموذجا من الصلصال .

١٧ - تكتب مقالا عن طريقة اقناع الناس وتوجيههم .

١٨ - تعمل ملقنا فى تمثيلية يقوم بها هواه .

١٩ - تحترف الطب .

٢٠ - تحترف النحت .

٢١ - تحترف الصحافة .

يعطى الاختبار عشرة مقاييس ميل بالاضافة الى مقياس تحقق Verification scale للكشف عن الاعمال ، أو سوء الفهم ، واختيار الاجابات المقبولة والمرغوب فيها اجتماعيا ولكنها غير متوقعة . وتتضمن مقاييس الميل العشرة الميل الآتية : الخلوى ، والميكانيكى ، والحسابى ، والعلمى ، والاقناعى ، والفنى ، والأدبى والموسيقى ، والعدالة الاجتماعية، والكتابى ، وقد وضع كيودر للاختبار معايير خاصة بكل من الذكور والاناث لتلاميذ الثانوى ، والجامعة ، والكبار . وتخطط الدرجات الكلية على مجالات الميل العشرة على قائمة معايير مثينة .

وتدور معاملات ثبات اختبار كيودر حول ٠.٩٠ ، كما يبدو الثبات باعادة الاختبار بعد عام أو اقل مرضيا ، أما بعد فترات أطول فليس هناك قدر كاف من المعلومات عنها ، وما هو موجود يشير الى أنه ، وخاصة حالة تلاميذ الثانوى ، كانت هناك تغيرات فى مجالات الميل المرتفعة والمنخفضة عندما كانت الفترة بين الاختبارات تمتد الى عدد من السنين (١٨ - ٤٧٥) ويبدو أن كلا من اختبار سسترونج وكيودر عرضة للكذب أو التحريف ، ويزيد هذا بالنسبة الى اختبار كيودر نظرا لطبيعة عناصره الأكثر وضوحا وشفافية .

أما دراسات صدق الاختبار فقد أجريت أساسا بمقارنة الاختبار بمحك الرضا فى العمل . ومن أكبر هذه الدراسات الطولية الشاملة تلك التى أجريت على ١١٦٤ تلميذا أعطوا اختبار كيودر فى الثانوى ، ثم أعطوا استبياننا

للرضا في العمل بعد عشر سنوات ، وعند ذاك كان عدد ٧٢٨ من هؤلاء ملتحقين بأعمال « متسقة » مع أنماط ميولهم الأصلية ، و ٤٣٦ في أعمال « غير متسقة » مع هذه الأنماط . وكانت نسبة العاملين الراضين في أعمالهم في الجماعة المتسقة هي ٦٢٪ ، بينما كانت النسبة هي ٢٤٪ في الجماعة غير المتسقة ، ومن ناحية أخرى كانت نسبة العاملين غير الراضين ٨٪ في الجماعة المتسقة ، و ٢٥٪ في الجماعة غير المتسقة (١٨) . وقد أيدت هذه النتائج دراسات أخرى في ذات الاتجاه .

وقد تطور حديثا عن اختبار كيودر اختبار آخر هو اختبار كيودر لمسح الميل العام (٤٠) Kuder General Interest Survey

وهو عبارة عن مراجعة وامتداد للاختبار الأصلي في الاتجاه الأدنى من الميول وقد صمم للأعمار من نحو ١٢ حتى ٢٤ ، ويستعمل لغة أبسط والمفاهيم أسهل من الاختبار السابق ، ولهذا فهو لا يتطلب مستوى رفيعا من القراءة . وقد وضعت معايير مئينية أمريكية لأكثر من عشرة آلاف ولد وبنت ، ولعينات قليلة من الكبار وما زال البحث جاريا للتحقق مما إذا كانت هذه الأداة تناسب الكبار كما تناسب من هم أصغر سنا ، وفي هذه الحالة قد ينتهي الأمر إلى أن يحل اختبار كيودر لمسح الميل العام محل الاختبار السابق . وتشير الدراسات التي أجريت على الاختبارين أن الارتباطات بينهما تعادل معاملات الثبات الخاصة بهما ، أي أنها تشير إلى درجة التشابه الكبير في قياس الميول عند الكبار .

وثمة صورة أخرى لاختبار كيودر للميول المهنية وهي (٤١) :
Kuder Occupational Interest Survey (KOIS)

اختبار كيودر لمسح الميل المهني ، وقد تطور هذا الاختبار وأنشئ على أساس المحك وبطريقة متشابهة لتلك التي اتبعتها سترونج في اختبارها ، كما سبق بيانه ، إلا أنه في هذه الصورة لم يستعمل جماعة الرجال بصفة عامة كما هو الحال عند سترونج ، وبدلا منها فإن درجة المفحوص على كل مقياس مهنة يعبر عنها ارتباط بين نمط ميوله ونمط ميول جماعة مهنية معينة . ويمكن الحصول حاليا على درجات لعدد ٧٩ مهنة و ٢٠ مجالا من مجالات الدراسة

الجامعية للذكور . و ٦٥ مهنة و ٢٥ مجال دراسة جامعية للاناث .
وتتنوع المهن التي يتناولها هذا الاختبار قنوعا كبيرا في مستواها فهي تتراوح
بين الخباز وسائق الشاحنة (سيارة النقل) من طرف ، والكيمائى والمحامى
من الطرف الآخر ، وقد ساعد حذف جماعة الرجال أو النساء بصفة عامة ،
أو جماعة المرجع ، على اتساع المجال الذى تستوعبه هذه الأداة الواحدة .
ويتضمن هذا الاختبار بالإضافة الى مقاييس المهنة ومقاييس مجالات الدراسة
الجامعية مقياس تحقق وثمانية مقاييس تجريبية تستعمل الآن لأغراض البحث
العلمى فقط .

وكما مر فى اختبارات كيودر كلها فان عناصر الاختبار مصاغة فى
صورة الاختيار الاجبارى ، ويغضى محتوى العناصر مجالات الميل المتضمنة
اختيار كيودر للميول المهنية مضافا اليها مجالات أخرى مثل تفضيل العمل
مستقلا ، وتجنب الصراع ، الميل للمواقف المألوفة والمستقرة .

وقد اثبت كيودر بتحليل احصائى موسع لثلاثة الاف مفحوص ان نظام
المقاييس المهنية المشتقة من جماعات المرجع الخاصة بالرجال بصفة عامة .
كما تظهر الدراسات التى أجريت على ثبات هذا الاختبار وصدقه أنها مرضية
الى حد بعيد ، وقد اعتمدت هذه الدراسات على ٢٠ مقياسا مهنيا ، ولا يزال
امام الاختبار ومقاييسه المهنية الأخرى مجال لدراسة صدقها وخاصة
بالنسبة الى محك خارجى ، (١٨ ، ٤١) ، ولن تستقر قيمة هذا الاختبار
النهائية الا بعد ان تتأكد بالبحث وبالتطبيق فى مجالات عملية متنوعة .

أما الصورة العربية لاختبار كيودر الميل المهنية ، والتى سبقت الإشارة
اليها ، فقد قام باعدادها وصياغتها احمد زكى صالح ، وتتكون من كراسة
للاختبار نفسه ، وورقة يضع المفحوص علامة x فى الخانة الأكثر تفضيلا
وأخرى فى الخانة الخاصة بأقل تفضيل لأساليب النشاط الثلاثة ، ثم تحسب
الدرجات الخام الخاصة بكل ميل بعد ان يتأكد المصحح من صدق الاجابة ،
ثم تحول الدرجات الخام الى مئينيات ويرسم تخطيط الميل على بطاقة تخطيط
خاصة بالبئين وأخرى خاصة بالبئات .

وللاختبار كراسة تعليمات تتضمن معلومات عن وصف الاختبار وهدفه

والميول التي يقيسها ، ومفاتيح التصحيح ومقياس الصدق ومفتاحه (ص) ، ثم وصف لاستخراج معايير الاختبار ، وبيانات عن ثباته وصدقته . وقد حظى هذا الاختبار باهتمام علمي واسع في مصر ، واستعمل ولا يزال في عدد من البحوث في مجال الميول المهنية (١ : ٢٦١ - ٢٦٤) .

ثالثا : اختبار لى - ثورب The Lee - Thorpe Inventory

بينما يمثل اختبار سترونج الأساس التجريبي في بناءه وتصحيحه ، ويمثل كيودر التنوع والتعدد ، نجد أن اختبار لى - ثورب يمثل اتجاها ثالثا في بناء اختبارات الميول . هذا الاختبار يتكون من مجموعة من الأمثلة التي اختيرت ونظمت على أساس الحكم وليس على الأساس الإحصائي . هذا الاتجاه الذي يمكن أن يسمى الاتجاه المنطقي ، مضاد لاتجاه سترونج الذي يعتمد على النتائج الإحصائية . ويشبه هذا الاتجاه المنطقي الأسلوب الذي يتبع في بناء اختبارات القدرة والاستعداد حيث تعين مجموعة كبيرة جدا من المواقف ثم تختار عنصر عشوائيا من هذه المجموعة الضخمة .

وقد بدأ لى ثورب بوصف المهن الموجودة في قاموس العناوين المهنية Dictionary of Occupational Titles الذي يصف ويصنف كل المهن والأعمال تقريبا في المجتمع الأمريكي ، واختيرت في ستة مجالات مهام تمثل ثلاثة مستويات للمستولية : مرتفع ، ومتوسط ، منخفض وتذكر أوصاف موجزة للأعمال في أزواج ، ويبين المفحوص ما يفضله من واحد من زوج ، وتمثل الدرجة التكرار النسبي للاختبارات في أقسام : الشخصية الاجتماعية ، والطبيعية ، والميكانيكية ، والأعمال ، والفنون ، والعلوم .

ويختلف هذا الاختبار عن اختبار سترونج في أن الاختبار الأصلي وتصنيف عناصره يعتمد كلية على أوصاف للمهنة وليس على دليل واقعي على أن أشخاصا في المهنة يحبون نشاطا معينا كما يختلف عن اختبار كيودر في أن التجميع الأول للعناصر بنى على أساس تحليل منطقي وليس على أساس العزل الإحصائي للعوامل وعلى الرغم من أن التصنيف الأصلي للعناصر اعتمد كلياً على أساس منطقي فقد أجريت دراسات ارتباطية تالية بهدف رفع

مستوى التجانس ، وعند مراجعة الاختيار استبعدت العناصر التي لا تشترك مع بقية العناصر في القسم الواحد (٢٧) .

الميل والشخصية : اتجاهات وتوقعات :

كانت الدراسات الأولى في الميل تعتمد اعتمادا كبيرا على المقارنة الواقعية الخالصة لجماعات مهنية على أساس عناصر مختارة بطسريقة عشوائية غالبا ، أى أنه لم يكن هناك نظرية خاصة بطبيعة الميل ، ولا بأنماط الميل التي يجب ان نوليها عناية خاصة ، وكذلك لم تكن هناك نظرية عن تكوين المهن والأعمال ، وكانت تفسيرات درجات اختبارات الميل تتم على أساس عملى غير نفسى (٢٧ : ٤٢٦) . وقد تغير الاتجاه في قياس الميل منذ بداية نشر سترونج اختبارده ، ويمكن القول ان اختبار سترونج واختبار كيورد أديا الى بداية نظرية في الميل . وقد وجد جيلفورد وزملاؤه (٢١) أكثر من عشرين عامل ميل يبدو أن معظمها يعكس أساليب شخصية عامة وليس اتجاهات مهنية . وأضاف جيلفورد الى ما هو معروف من ميل بمثل الميكانيكية والعلمية وغيرها ، وأضاف المغامرة في مقابل الأمن ، والتذوق الجمالى ، والمسيرة الثقافية أو النظام ، والعدوان وأبعاد ميل أخرى كثيرة وكذلك أدت الدراسات الطولية ودراسات الحالة الى اعطاء صورة أكثر وضوحا عن أهمية الميل ومفزاها . أى ان هناك تطورا ونموا متوقعا في اتجاه نظرية متكاملة عن الميل .

وثمة اتجاه آخر نام أيضا هو التفسير المنظم للميل في ضوء مفاهيم أساسية للشخصية . فبينما اتى على الميل حين من الدهر كان ينظر اليها هيه على أنها عمليات اشتراط تقوم على الصدفة . نجد ان الميل الآن ينظر اليها على أنها تعبير عن حاجات عميقة وعن أنماط توافق معينة . وربما تكون اختبارات الميل أكثر كفاءة من عدد كبير من الأساليب الأخرى في قياس الشخصية نظرا لتنوع محتواها وزيادة تقبل الشخص لها .

وقد كشف عدد كبير من الدراسات عن ارتباطات ذات دلالة بين الميل المهنية المداسة وجوانب أخرى للشخصية . فمن هذه الدراسات ما أظهر مثلا انه في حالات ادمان الخمر ، والجنسية المثلية ، والعصاب ، وغيرها من

نواحى العجز الذهانية كانت تسود كل جماعة من هذه الجماعات انماط ميول مهنية خاصة مميزة لها (٢٥ ، ٢٨) . وبينت دراسات أخرى أجريت على تلاميذ المدارس والجامعات والعمال الصناعيين والمرضى الراشدين أن الاستهداف للحوادث يرتبط ارتباطاً ذا دلالة بمقاييس معينة على اختبار سترونج ، وكذلك ثبت أن هناك ارتباطاً بين درجات اختبار سترونج وكيودر واختبارات الشخصية مثل اختبار مينسوتا المتعدد الأوجه وغيره (١٨ : ٤٩١)

ومن جهة أخرى فإنه من المعترف به الآن أن اختبار مهنة غالبا ما يعكس حاجات الشخص العاطفية الأساسية ، وأن التوافق المهني يمثل جانبا أساسيا في توافق الشخص العام في الحياة . ذلك لأن الشخص عندما يختار مهنة ، فإنه إلى حد بعيد ، يختار أساليب التوافق وأنماط الحياة والأدوار التي يحس أنها قريبة إليه ومناسبة له . وعلى هذا فإن قياس الميول المهنية ، وبصفة خاصة تحديد الجماعات المهنية التي يشارك الفرد أعضائها ميولهم واتجاهاتهم ، يصبح أمرا هاما في فهم مختلف الشخصيات . ومما يؤيد التأكيد الحالي لاتخاذ الاختبارات المهنية مؤشرات للشخصية ذلك الاختبار الذي وضعه هولاند (٢٠) Holland Vocational Preference Inventory وقد استعمل فيه العناوين المهنية لقياس خصائص الشخصية . وما على المفحوص في هذا الاختبار إلا أن يوضح ما إذا كان يحب أو يكره كل واحدة من ثلاثمائة مهنة . وعلى الرغم من أن هذا الاختبار قد اعتمد اعتمادا كبيرا على صدق المحتوى الداخلي ، فإن مقاييسه قد أظهرت اختلافات دالة بين عينتين متناظرتين أحدهما تتكون من أفراد مصابين بأمراض نفسية ، وتتكون الأخرى من أفراد عاديين ويعطى الاختبار عشرة مقاييس للشخصية ، منها النشاط البدني والنزعة العقلية ، والمسؤولية ، والمسايرة وما زالت قيمة هذا الاختبار رهن البحث والاستعمال العملي إلا أنه يمثل اتجاها جديدا في دراسة الشخصية عن طريق الميول .

من هذا يتضح أنه من المتوقع أن تتطور نظرية خاصة بالميل ، وأن يمثل قياس الميول مجالا حيا في تطور الاختبارات وفي البحث العلمي ، وأن تنمو وتتوسع الدراسات حول العلاقة بين الميول والشخصية ، وأن يتجه تفسير اختبارات الميول إلى مزيد من العمق والدقة .

مراجع الباب السادس

- ١ - أحمد زكى صالح : علم النفس التربوى . القاهرة : مكتبة النهضة المصرية (ط ١٠) ، ١٩٧٢ .
- ٢ - أحمد زكى صالح : الأسس النفسية للتعليم الثانوى . القاهرة : مكتبة النهضة المصرية (ط ٢) ، ١٩٧٢ .
- ٣ - أحمد عبد العزيز سلامة : تطبيق اختبار تفهم الموضوع على حالات مصرية . رسالة ماجستير ، كلية التربية جامعة عين شمس ، ١٩٦٢ .
- ٤ - سيد أحمد عثمان : المسئولية الاجتماعية . القاهرة : مكتبة الأنجلو المصرية ، ١٩٧٢ .
- ٥ - سيد أحمد عثمان : علم النفس الاجتماعى التربوى . القاهرة : مكتبة الأنجلو المصرية ، ١٩٧٠ .
- ٦ - سيد أحمد عثمان ، فؤاد أبو حطب : التفكير ، دراسات نفسية ، القاهرة . مكتبة الأنجلو المصرية ، (ط ٢) ١٩٧٧ .
- ٧ - سيد محمد غنيم : مدى صلاحية اختبار بقع الحبر لروشاش لقياس الذكاء . رسالة ماجستير ، كلية التربية جامعة عين شمس ، ١٩٥٥ .
- ٨ - سيد محمد غنيم ، هدى عبد الحميد برادة : التشخيص النفسى : دراسات فى اختبار الروشاش ، القاهرة : دار النهضة العربية (ح١) ، ١٩٦٥ ، (ح ٢) ١٩٦٦ .
- ٩ - سيد محمد غنيم ، هدى عبد الحميد برادة : الاختبارات الإسقاطية . القاهرة ، دار النهضة العربية ، ١٩٦٤ .
- ١٠ - فاروق إبراهيم أبو عوف : دراسة مقارنة لمعرفة مدى الافادة من اختبار رورشاخ (بقع الحبر) للتمييز بين حالات العصاب القهرى

والأسوياء . رسالة ماجستير ، كلية التربية جامعة عين شمس ،
١٩٧٢ .

١١ - فاروق ابراهيم أبو عوف : مدى صلاحية اختبار رورشاخ لقياس
الابتكارية . رسالة دكتوراه ، كلية التربية جامعة عين شمس ،
١٩٧٦ .

١٢ - فاروق ابراهيم أبو عوف : دراسة مقارنة لبعض سمات الشخصية
بين القادة وغير القادة باستخدام اختبار بقع الحبر لرورشاخ .
الكتاب السنوى فى علم النفس ، الجمعية المصرية للدراسات النفسية،
المجلد الرابع ، القاهرة . مكتبة الأنجلو المصرية ، ١٩٨٥ .

١٣ - فؤاد أبو حطب : التحليل العلمى للسلوك الخلقى . الكتاب السنوى
فى التربية وعلم النفس (تحرير سعيد اسماعيل على) . المجلد
الأول ، القاهرة : عالم الكتب ، ١٩٧٣ .

١٤ - فؤاد أبو حطب : التفضيل الفنى وسمات الشخصية . المجلة
الاجتماعية القومية ، يناير ١٩٧٣ .

١٥ - لويس كامل مليكه ، محمد عماد الدين اسماعيل : الشخصية وقياسها .
مكتبة النهضة المصرية ، ١٩٥٩ .

١٦ - نجيب اسكندر ابراهيم ، لويس كامل مليكه ، رشدى فام منصور :
الدراسة العلمية للسلوك الاجتماعى . القاهرة : مؤسسة المطبوعات
الحديثة ، ١٩٦٦ .

١٧ - يوسف الشيخ ، جابر عبد الحميد جابر : سيكولوجية الفروق الفردية
القاهرة : دار النهضة العربية ، ١٩٦٤ .

18. Anastasi, A. Psychological Testing. New York, Macmillan,
1982.

19. Atkinson, J.W. (ed.) *Motivation, fantasy action, and society.* Princeton, N.J. Van Nostland, 1958.
20. Atkinson, J.W. and Feather, N.T. *A theory of achievement Motivation.* New York : Wiley, 1966.
21. Barkes, R.G., Kounin, G. and Wright, H.F. (eds.) *Child behaviour and development.* New York : McGraw-Hill, 1943.
22. Beck, S.L. *Rorchach's test. I : Basic processes,* New York : Grune and Stratton, 1944.
23. Beck, S.L. *Rorchach's test. III : Advances in interpretation.* New York : Grune and Stratton, 1952.
24. Bell, H.M. *The Adjustment Inventory.* Palto Alto, Calif : Consulting Psychologists, 1934-1939.
25. Carnes, G.D. Vocational interest characteristic of abnormal personalities. *J. of Counsel, Psychol.* 1964, II, 272-279.
26. Combs. A.W., a Snygg, D. *Individual Behaviour,* New York : Harper, 1959.
27. Cronbach, G.L. *Essentials of Psychological Testing,* New York : Harper, 1984.
28. Drasgow, J. and Carkhuff, R.R. Kuder neuro-psychiatric Keys before and after psychotherapy. *J. of Counsel. Psychol.* 1964, 11, 67-71.
29. Flanagan, G.C. *Factor analysis in the study of personality.* Stanford : Stanford Univ. Press, 1935.
30. Goldner, R.H. Individual differences in whole-part approach and flexibility-rigidity in problem solving. *Psycholo. Monogr.* 1957, 71, No. 21.
31. Guilford, L.P. et al., A factor analysis study of human interests. *Psychol. Morograph,* 1954, 68, No. 41.
32. Hartshorne, H. and May, *Studies in deceit.* New York : Macmillan, 1928.
33. Hartshorne, H. and May, M. *Studies in service and self-control,* New York : Macmillan, 1929.
34. Hartshorne, H. and May, M. *Studies in the organization of character,* New York : Macmillan, 1930.

35. Hathaway, S.R., and Mckinley, G.C. Minnesota Multiphasic Personality Inventory : Manual for administration and scoring. New York : Psychological Corporation, 1967.
36. Holland, J.L. A personality Inventory employing occupational titles. J. of Appl. Psychol., 1958, 42, 336-342.
37. Holtzmann, W.H., et al. Inkblot perception and personality- Holtzmann Inkblot Technique. Austin, Texas : Univ. of Texas Press, 1961.
38. Klein, G.S. and Schlesinger, H.L. Perceptual attitudes toward instability : I. Prediction of apparent movement experiences from Research response. J. Personol., 1951, 19, 289-302.
39. Krech, D., Crutchfield, R.S. and Ballachey, E. L. Individual In Society. New York : McGraw-Hill, 1962.
40. Kuder, G.F. General Interest Survey, Manual Chicago : Science Research Associates, 1964.
41. Kuder, G.F. Kuder Occupational Interest Survey : General Manual. Chicago : Science Research Associates, 1966.
42. Kunce, J.T. Vocational Interests and accident proneness. J. of Appl. Psychol., 1967, 51, 223.
43. Lindzey, G. and Borgatta, E.F. Sociometric measurement. In Lindzey, G. (ed.) Handbook of Social Psychology. Cambridge, Mass. Addison-Wesley, 1954, Vol. I, Ch. II.
44. Luchins, A.S. Mechanization in Problem solving. Psychol. Monogr., 1942, 54, No. 6.
45. Masling, J. The influence of situational and interpersonal variables in projective testing. Psychol. Bulletin, 1960, 56, 65-85.
46. Masling, J. Differential introduction of examiners and Rorschach responses. J. of Consult. Psychol., 1965, 29, 198-201.
47. McClelland, D.C., et al. The achievement motive. New York: Appleton-Century, 1953.
48. McClelland, D.C. The achieving society. New York : Free Press, 1961.
49. Moreno, G.L. Who Shall Survive ? Washington : Nervous and Mental Diseases Publishing Co., 1934.

50. Murray, H.A., et al. Explorations in personality. New York: Oxford Univ. Press, 1938.
51. Myers, I.B. The Myers-Briggs Type Indicator : Manual Princeton, N.J. : Educational Testing Service, 1962.
52. Peters, H.N. and Campbell, G.T. Diagnosis of training needs of B-29 mechanics from supervisory ratings and self-report. AFPTRC Tech. Memorandum 55-12, 1955.
53. Roe, A. Two Rorchach Scoring Techniques : The inception Technique and the basic Rorchach. J. Abnorm. Soc. Psychol., 1952, 47, 263-264.
54. Rogers, C.R. Counseling and Psychotherapy, Cambridge, Mass; Riverside, 1942. -
55. Rogers, C.R. Client-Centered Therapy. Cambridge, Mass. : Riverside, 1951.
56. Ruggles, R. and Allport, G.W. Recent Application of the A.-S. Reaction Study. J. Abnorm. Soc. Psychol., 1934, 34, 518-528.
57. Sarason, S.B. The clinical interaction with special reference to the Rorchach. New York : Harper, 1954.
58. Strong, E.K. and Campbell, D.P. Manual for strong Vocational Interest Blank Stanford, Calif : Stanford Univ. Press, 1966.
59. Taylor, Janet A.A. personality scale of manifest Anxiety. J. Abnorm. Soc. Psychol., 1953, 48, 285-290.
60. Wallen, R.W. Clinical Psychology. New York : McGraw-Hill, 1956.

خاتمة

تقويم أدوات التقويم

عرضنا في فصول هذا الكتاب المفاهيم الأساسية التي تستخدم في الحكم على صلاحية الاختبارات وأدوات التقويم بوجه عام ، ثم طبقنا هذه المفاهيم على أمثلة رئيسية من اختبارات الذكاء والقدرات والاستعدادات والتحصيل والشخصية والميول .

ونعرض في هذه الخاتمة نموذجا يصلح لتقويم الاختبارات وأدوات التقويم الأخرى يتألف من ٢٨ بنداً ، وفيه تتطلب البنود رقم ١ إلى رقم ٩ وصفا بسيطا للاختبار يمكن أن نحصل عليه عادة من كراسة تعليماته والدليل الخاص به ، ويتطلب البند رقم ٦ وصف الطابع العام للاختبار . كما أن البند رقم ٧ (تاريخ النشر) ليس على درجة كبيرة من الأهمية لأن بعض الاختبارات القديمة قد تكون ممتازة . وإى اختبار قديم يجب أن يفحص بعناية ودقة ، فقد تكون بعض أسئلته غير ملائمة لتغير العصر ، وقد تكون معايير غير مفيدة ، وقد يكون دليله غير كامل . وفى هذا البند يجب أن تسجل أى طبعات جديدة أو أى تعديلات طرأت على الاختبارات .

ويحتوى النموذج المشار إليه على البند رقم ١١ الذى يتطلب وصف الاختبار من حيث عناصر وأسئلته وتصحيحه من أجل تكوين انطباع عام عن الاختبار ولا شك أن هذا البند هو الذى يحدد مدى جاذبية الاختبار للمفحوص والمرجل العادى . أما البند ١٢ يتطلب تحديد مقاصد المؤلف سواء كانت انتقاء أو توجيها أو علاجاً أو تقويماً مدرسياً ، وبالقائى يحدد بهدف المؤلف إلى قياسه بالاختبار . وقد تتضمن بعض كراسات التعليمات الأسلوب الإحصائى الذى استخدمه المؤلف فى انتقاء الأسئلة أو تحليلها وحساب صدق تلك الأسئلة .

ويتضمن البند ١٣ حكماً على تعليمات الاختبار فى ضوء الأسس العامة التى عرضناها فى الفصل الثانى من الكتاب ، أما البند ١٤ فيتطلب أن نقوم بتحليل نظرى لأسئلة الاختبار للحكم على ما تتضمنه من قدرات أو خبرات أو عادات أو سمات شخصية ، وكيف تؤثر هذه العوامل كلها أو بعضها فى

درجة الاختبار . هذا التحليل ضرورى لكل درجة من درجات الاختبارات الفرعية التى يحتوئها الاختبار وبالطبع نحن لا نستطيع أن نحدد جميع العوامل المسهمة فى درجة الاختبار . ولكن مثل هذا العمل التحليلى بشير عددا من الأسئلة يمكن أن تستخدم فى تفسير نتائج الاختبار . ويجب أن يسجل التقرير ما تدل عليه الدرجة فى الاختبار سواء كان حكماً هذا يتفق مع مقصد المؤلف أو لا يتفق ، كما يوضح العوامل غير المرتبطة التى تؤثر فى الدرجة وقد تفسدها . ويمكن للباحث أن يلجأ الى تطبيق الاختبار على نفسه أو على مفحوص آخر لديه خبرة بالقياس النفسى ثم يلاحظ الباحث أداءه فى الاختبار ، أما فى حالة اختبار المفحوص لنفسه فإن الملاحظة تصبح من نوع التأمل الباطنى introspection ، ورغم الاعتراضات المنهجية على هذه الطريقة إلا أنها تفيد كثيراً فى إعطاء كثير من الدلالات عن طبيعة الاختبار .

وتدل البنود من ١٦ - ١٩ ما يتوفر عن صدق الاختبار من بيانات تجريبية ، وذلك للحكم على مدى صلاحية هذه البيانات لأغراض الاختبار ، أما البنود من ٢٠ - ٢٢ فتشمل بيانات ثبات الاختبار ، ويتصل البند ٢٣ بثبات الاختبار من حيث فاحصه ومصحيبه ، والمعايير من حيث دقة تفسير الأداء فى الاختبار كما يدل عليها البنود ٢٤ ، ٢٥ .

ويشير البند ٢٦ الى تلخيص العروض الناقدة للاختبار فى المؤلفات والمجلات المتخصصة . أما البند ٢٧ وموضوعه « تقويم عام للاختبار » فيتضمن تلخيصاً عاماً ونهائياً لنواحي الضعف والقرّة مع اعتبار النواحي العلمية والفنية جميعاً ، ومن الملائم مقارنة الاختبار بغيره من الاختبارات التى تقيس نفس الوظائف العامة ، بالإضافة الى تسجيل الملاحظات العامة للفاحصين والمفحوصين وتعليقاتهم الخاصة عليه .

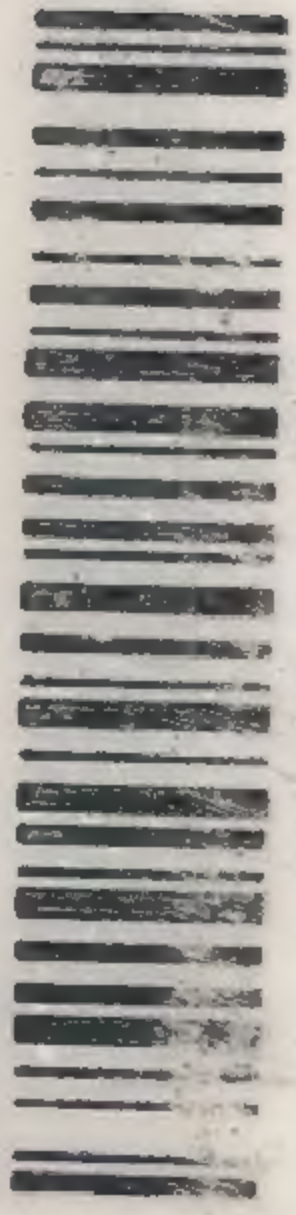
وهكذا يتوافر للفاحص قائمة تحليلية للاختبار بحيث يستطيع الفاحصون - وخاصة أولئك المشاركون فى البرامج الشاملة للتقويم سواء التربية أو الصناعة أو القوات المسلحة - أن يختاروا الاختبار المناسب للغرض المناسب فى الوقت المناسب ، وبأقل فاقد ممكن فى الجهد والوقت والمال .

نموذج لتقويم الاختبارات

- ١ - عنوان الاختبار .
- ٢ - اسم مؤلف الاختبار ومترجمه (ان وجد) .
- ٣ - الناشر .
- ٤ - صور الاختبار والجماعات التي يمكن أن تطبق عليها .
- ٥ - المعالم العملية (اية خصائص عملية تتصل بالأداء فيه او تصحيحه) .
- ٦ - نوع الاختبار (اختبار قدرات او شخصية مثلا) .
- ٧ - تاريخ النشر .
- ٨ - التكاليف ، كراسة التعليمات والأسئلة ، ورق الاجابة .
- ٩ - الزمن المطلوب .
- ١٠ - الغرض الذي من أجله نحكم على الاختبار .
- ١١ - وصف الاختبار من حيث عناصره ، وتصحيحه .
- ١٢ - غرض المؤلف وأساس اختبار عناصر الاختبار عنده .
- ١٣ - دقة التعليمات ووضوحها ، والتدريب المطلوب لمن يقوم بعملية اعطاء الاختبار ، ودرجة تقنينها .
- ١٤ - الوظائف أو السمات العقلية الممثلة في كل درجة .
- ١٥ - تعليقات وملاحظات عن تصميم الاختبار وبنائه (ورقة الأسئلة وطريقة الاجابة) .
- ١٦ - الصدق التنبؤي (المحك ، وعدد ونوع الحالات التي استخدمت في حسابه ، والنتيجة) .
- ١٧ - الصدق التلازمي (المحك ، وعدد ونوع الحالات المستخدمة ، والنتيجة) .
- ١٨ - اية أدلة تجريبية أخرى تبين ما يقيسه الاختبار (صدق التكوين الفرضي) .
- ١٩ - تعليقات وملاحظات عن صدق الاختبار بالنسبة لأغراض معينة سيسعمل من أجلها الاختبار .
- ٢٠ - ثبات الاختبار بحساب معامل الاستقرار (الطريقة التي اتبعت للوصول اليه وعدد ونوع الحالات ، والنتيجة) .

- ٢١ - ثبات الاختبار بحساب معامل التكافؤ (الطريقة التي اتبعت للوصول إليه ، وعدد نوع الحالات ، والنتيجة) .
- ٢٢ - ثبات الاختبار بحساب معاملات الاتساق الداخلى أو اتساق الأسئلة (الطريقة التي اتبعت للوصول اليها ، وعدد ونوع الحالات ، والنتيجة) .
- ٢٣ - موضوعية الاختبار من حيث التصحيح تفسير الدرجات .
- ٢٤ - المعايير (نوعها ، وعدد ونوع الحالات) .
- ٢٥ - تعليقات وملاحظات عن دقة فئات الاختبار وصحة المعايير للغرض الذى سيستعمل الاختبار من أجله .
- ٢٦ - تعليقات وملاحظات من قاموا باستعراض للاختبار وتقويمه ونقده .
- ٢٧ - تقويم عام للاختبار .
- ٢٨ - المراجع .

Bibliotheca Alexandrina



0658938